

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 18:17:26 ; Search time 400 Seconds
(without alignments)
6251.504 Million cell updates/sec

Title: US-09-386-850-7_COPY_1_510

Perfect score: 510
Sequence: 1 AACTGACTGACGAGAGC.....TTTTCAGGTGGTGCAMA 510

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445938

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	658	10	US-09-386-850-7
2	65	12.7	65	10	US-09-908-975-3241
3	53.4	10.5	153	9	US-09-919-580-624
4	39.4	7.7	65	10	US-09-908-975-29028
5	39.4	7.7	56153	13	US-10-221-714A-519
6	38.2	7.5	1391	17	US-10-713-836-5
7	38.2	7.5	6022	13	US-10-221-714A-383
8	38.2	7.5	40304	12	US-09-997-722-280
9	37.4	7.3	19653	15	US-10-311-455-1308
10	36.8	7.2	630	17	US-10-437-963-78930
11	36.6	7.2	367378	15	US-10-312-841-2
12	36.4	7.1	5127	15	US-10-239-676-132
13	36.4	7.1	6222	15	US-10-311-455-666
14	36.2	7.1	777	13	US-10-027-632-135035

15	36.2	7.1	777	13	US-10-027-632-135036	Sequence 135036,
16	36.2	7.1	777	13	US-10-027-632-135037	Sequence 135037,
17	36.2	7.1	777	16	US-10-027-632-135035	Sequence 135035,
18	36.2	7.1	777	16	US-10-027-632-135036	Sequence 135036,
19	36.2	7.1	777	16	US-10-027-632-135037	Sequence 135037,
20	36.2	7.1	5815	17	US-10-257-166-105	Sequence 105, App
21	36.2	7.1	8776	17	US-10-257-166-150	Sequence 150, App
22	36.2	7.1	659158	9	US-09-771-208-20	Sequence 20, Appl
23	36	7.1	644	9	US-09-759-143-752	Sequence 752, App
24	36	7.1	644	9	US-09-780-669-752	Sequence 752, App
25	36	7.1	644	9	US-09-822-827-752	Sequence 752, App
26	36	7.1	644	9	US-09-895-793-752	Sequence 752, App
27	36	7.1	644	9	US-09-895-814-752	Sequence 752, App
28	36	7.1	644	14	US-10-012-896-752	Sequence 752, App
29	36	7.1	644	15	US-10-144-678A-752	Sequence 752, App
30	36	7.1	644	15	US-10-294-025-752	Sequence 752, App
31	35.8	7.0	7029	15	US-10-311-455-593	Sequence 593, App
32	35.6	7.0	549	13	US-10-621-901-266	Sequence 266, App
33	35.6	7.0	4255	15	US-10-311-455-2006	Sequence 2006, Ap
34	35.6	7.0	367378	15	US-10-312-841-1	Sequence 1, Appli
35	35	6.9	14253	15	US-10-311-455-1467	Sequence 1467, Ap
36	34.8	6.8	24259	13	US-10-221-714A-415	Sequence 415, App
37	34.6	6.8	65	10	US-09-908-975-27963	Sequence 27963, A
38	34.4	6.7	523	9	US-09-864-761-30606	Sequence 30606, A
39	34.4	6.7	1721	13	US-10-027-632-98304	Sequence 98304, A
40	34.4	6.7	1721	16	US-10-027-632-98304	Sequence 98304, A
41	34.4	6.7	10039	15	US-10-311-455-2015	Sequence 2015, Ap
42	34.4	6.7	33578	16	US-10-085-117-238	Sequence 238, App
43	34.2	6.7	1150	13	US-10-115-123-69	Sequence 69, Appl
44	34.2	6.7	1150	15	US-10-012-542-69	Sequence 69, Appl
45	34.2	6.7	1310	17	US-10-437-963-60606	Sequence 60606, A

ALIGNMENTS

RESULT 1
US-09-386-850-7

; Sequence 7, Application US/09386850
; Publication No. US20030170620A1

; GENERAL INFORMATION:

; APPLICANT: Rosinski-Chupin, Isabelle R.

; Tronik, Diana

; Rougeon, Francois

; Seidah, Nabil

; TITLE OF INVENTION: Peptides and Polypeptides Derived

; from the Submaxillary Gland of the Rat, Corresponding

; Antibodies, Corresponding Hybridomas, and Uses of

; These Products for Diagnosis, Detection, or

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker & Mathis

; STREET: P.O. Box 1404

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-1404

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/386,850

; FILING DATE: 31-Aug-1999

; CLASSIFICATION DATA:

; APPLICATION NUMBER: 530

; APPLICATION NUMBER: US/08/476,120

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/153,277

; FILING DATE: 17-NOV-1993

; APPLICATION NUMBER: US 07/499,276

; FILING DATE: 19-JUL-1990

```
; APPLICATION NUMBER: PCT/FR88/00523
; FILING DATE: 11-OCT-1989
; APPLICATION NUMBER: FR 88/13353
; FILING DATE: 11-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 004900-129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-8620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 658 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: No. US20030170620A1 Relevant
; TOPOLOGY: No. US20030170620A1 Relevant
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Rat
; TISSUE TYPE: Submaxillary Gland
; CELL TYPE: Glandular
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: SMR1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..510
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-386-850-7
```

```
Query Match 100.0%; Score 510; DB 10; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.1e-151;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAGCTGACGACAGAGAGCTTCTGACGACGACATTTCCCGCTCAGAGATTCTTCCAAG 60
DB 1 AAACAGCTGACGACAGAGAGCTTCTGACGACGACATTTCCCGCTCAGAGATTCTTCCAAG 60

QY 61 GGGCTACCAAGATGAAGTACATGTTTATCTTTGGCTGTGGATCTCTTAGCATGC 120
DB 61 GGGCTACCAAGATGAAGTACATGTTTATCTTTGGCTGTGGATCTCTTAGCATGC 120

QY 121 TTCCAGTTCAGGTGAGGTGTGACAGGCCCAAGAGACAAATATCTTAGAGACAA 180
DB 121 TTCCAGTTCAGGTGAGGTGTGACAGGCCCAAGAGACAAATATCTTAGAGACAA 180

QY 181 GATCCTTCAACTCTTCTCTATCTTGTCTTCAGCCTGATCCCAATGGTGACAAATA 240
DB 181 GATCCTTCAACTCTTCTCTATCTTGTCTTCAGCCTGATCCCAATGGTGACAAATA 240

QY 241 GGAGTAACATCACTATACCTTAAATCTTCAACCACTGCTGTCTTGTAAATCTTCCC 300
DB 241 GGAGTAACATCACTATACCTTAAATCTTCAACCACTGCTGTCTTGTAAATCTTCCC 300

QY 301 GGTGTTTATCACTCGACACCAATGGTGTGTAACAAGGTACCACTGAATATCAATATCAGTGG 360
DB 301 GGTGTTTATCACTCGACACCAATGGTGTGTAACAAGGTACCACTGAATATCAATATCAGTGG 360

QY 361 CAGCTAACTGCTCCAGACCTACACCTCTAGCAATCTTCACTCACTCACTCACTCACTCA 420
DB 361 CAGCTAACTGCTCCAGACCTACACCTCTAGCAATCTTCACTCACTCACTCACTCACTCA 420

QY 421 GAACAAGCAATAACAAAAACAGATGCGAATAATCTCCAACACTACTGCGACTACCCAAAAT 480
DB 421 GAACAAGCAATAACAAAAACAGATGCGAATAATCTCCAACACTACTGCGACTACCCAAAAT 480

QY 481 TCCACTGATATTTTGAAGTGTGGCAAA 510
DB 481 TCCACTGATATTTTGAAGTGTGGCAAA 510
```

RESULT 2
US-09-908-975-3241

```
; Sequence 3241, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36888-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3241
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-908-975-3241
```

Query Match 12.7%; Score 65; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 277 CTCGTGTTCTTGTAAATCTTCCCGTTTATCACTGGACCAACATTTGTGTACAAGGT 336
DB 1 CTCGTGTTCTTGTAAATCTTCCCGTTTATCACTGGACCAACATTTGTGTACAAGGT 60

QY 337 ACCAC 341
DB 61 ACCAC 65
```

RESULT 3

```
US-09-919-580-624
; Sequence 624, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 624
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-919-580-624
```

Query Match 10.5%; Score 53.4; DB 9; Length 153;
Best Local Similarity 63.8%; Pred. No. 1.8e-06;
Matches 81; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
QY 69 AAAGATGAAGTCACTGATTTTGTATCTTTGGCCCTGTGGATCTTCTAGCATGCTTCCAGTC 128
DB 26 ARGATGAATCACTGATTTTGTATCTTTGGCCCTGTGGATCTTCTAGCATGCTTCCAGTC 85

QY 129 AGGTAGGGTGTGAGAGGCCCAAGAGACAAATATCTAGAGACAAACAGATCTTC 188
DB 86 TGGTGAAGTCAAGAGGCCCAAGAGACAAATATCTAGAGACAAACAGATCTTC 145

QY 189 AACTCTT 195
```

```
Db      146 ACCTTTT 152

RESULT 4
US-09-308-975-29028
; Sequence 29028, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29028
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-308-975-29028

Query Match      7.7%; Score 39.4; DB 10; Length 65;
Best Local Similarity 75.4%; Pred. No. 0.031;
Matches 49; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 424 CAAGCAATAACAAAACAGATGCAAAATCTCCAAACACTACTGCGACTACCCAAAATTCC 483
Db 1 CAAGCAACTACAATGCCAGCTGCCAGTATCTCCATACTACTCTCTACTGCTAGAGATTCC 60

QY 484 ACTGA 488
Db 61 ACTGA 65

RESULT 5
US-10-221-714A-519/c
; Sequence 519, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 519
; LENGTH: 56153
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-519

Query Match      7.7%; Score 39.4; DB 13; Length 56153;
Best Local Similarity 50.8%; Pred. No. 1.9;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 149 CAAGAAGACAAACATAATCTCTAGAAGACAACAAGATCTTCAACTCTTCCTCATTTCTTG 208
Db 11877 CAACATCTAACAAAATATACACGAAAAAACAACCTTAAAAAACCCTTAACACATAT 11818

QY 209 GTCTTCAGCCTGATCCCAATGGTGGCAATAGAGAGTAACAATCACATATACCTTAAATC 268
Db 11817 AACACTAAACTAATACTAATAATAAAAAAACAACATATCTCGAATTCACAAAAAC 11758

QY 269 TTCAACCCACCTCGTGTCTTGTAACTCTCCCGGTTTATCACTGGACCACTTGGTTG 328
Db 11757 ATCCCTCTCTCGAAAACTATTCAATCCCTATATTTTCTAACTCTTTCCTACTAATCATCG 11698

QY 329 TACAA 333
Db 11697 TACCA 11693

RESULT 6
US-10-713-836-5/c
; Sequence 5, Application US/10713836
; Publication No. US20040143871A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Isolation of Genes for Galactomannan
; TITLE OF INVENTION: Formation from Guar Seeds
; FILE REFERENCE: 1520
; CURRENT APPLICATION NUMBER: US/10/713,836
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/426,127
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/490,022
; PRIOR FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Cyamopsis tetragonoloba
US-10-713-836-5

Query Match      7.5%; Score 38.2; DB 17; Length 1391;
Best Local Similarity 58.3%; Pred. No. 0.48;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 189 AACTTCTCTCATTTATCTTGGTCTTCAGCTGATCCCAATGGTGGACAATAAGAGCTAAC 248
Db 1018 ACCTCTCGTTATGTTCTCTGAGCTTTCACGATCTCCAGCCAGTACCCCTTCAAGTAGT 959

QY 249 AATCACTATACCCCTTAAATCTTCAACCACTCGTGTCTTGTAAATCTTCCCGGT 303
Db 958 ACTCGCTCTCCAAGAAGATCTTCTCCCTCCAGCTGTCTTATATATCCGTCGGAT 904

RESULT 7
US-10-221-714A-383/c
; Sequence 383, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
```



```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78930
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7868C.1
US-10-437-963-78930

Query Match          7.2%; Score 36.8; DB 17; Length 630;
Best Local Similarity 48.6%; Pred. No. 0.83;
Matches 101; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 157 CRACATAATCTAGAGACAGACAGATCTTCACTCTCTCTCAATATCTGTGCTTCAG 216
Db 137 CATCATATCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 196

QY 217 CTGTATCCCAATGGTGACAAATAGGAGTAACAATCACTATACCTTTAAATCTTCAACCA 276
Db 197 CGTCATCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 256

QY 277 CTCCTGTTCTTGTATCTTCCCGTTTATCACTGGACCACTTGGTTGTACAAGGT 336
Db 257 CTCCTCTCTGTCATCATCTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 316

QY 337 ACCACTGAATATCAATATCAGTGGCAGC 364
Db 317 CACTCTCCATTGTCATCTCTATCAGCCGC 344

RESULT 11
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match          7.2%; Score 36.6; DB 15; Length 3673778;
Best Local Similarity 50.3%; Pred. No. 1.7e+02;
Matches 90; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 303 TTTTATCACTGGCCACCATTGGTTGTACAGTACACTGAATATCAATATCAGTGGCA 362
Db 308899 TTTTATTTTCAAAAACCGCAATTTATAACTTAACCAATTTTACCAAACTAACTTAAACA 308840

QY 363 GCTAATCTCTCCAGACCTTACACCTCTAAGCAATCTCTCTCTCTCTCTCTCTCTCTCT 422
Db 308839 ACNAATCTTAAATCTTAAACCCACATTTTATCTTATTAACCCCTTTTAAACCAAA 308780

QY 423 ACAAGCAATATCAAAAACAGATGCCAAAATCTCCAACTACTCGCACTACCCAAAATT 481
Db 308779 TAACACAAAAAACCAATTCTATCAACACAAATATACACCAATTTTACTACAACTTAAAT 308721

RESULT 12
US-10-239-676-132/c
; Sequence 132, Application US/10239676
; Publication No. US20030082609A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 132
; LENGTH: 5127
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (1976, 1981...1982, 1984, 1989...1990, 1995...1996, 5032, 5059)
US-10-239-676-132

Query Match          7.1%; Score 36.4; DB 15; Length 5127;
Best Local Similarity 52.0%; Pred. No. 4;
Matches 106; Conservative 0; Mismatches 96; Indels 2; Gaps 1;

QY 257 TACCCCTTAAATCTTCAACCACTCGTGTCTTCTTAACTCTCCGGTCTTATCACTGGAC 316
Db 2577 TACGATTTTCTCTAATCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAACA 2518

QY 317 CACCATTTGTTGTACAAAGTACCACTGAATATCAATATCAATATCACTGAGTGGCAGTAACTGCTCCAG 376
Db 2517 AAACAATAACAAAAAACCTTAACTTTATAAAAAAACAACAAACCCCTCTCTCCAA 2458

QY 377 ACCTACACCTCTAAGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 436
Db 2457 ACCCCCAAACTCTAACCCCAATTTCTCT--CCAACCTTAACCTACCCGAATAAAAAATATAA 2400

QY 437 AAACAGATGCCAAAATCTCCAACA 460
Db 2399 AATAACACCAAAATAACCAATA 2376

RESULT 13
US-10-311-455-666/c
; Sequence 666, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 666
; LENGTH: 6222
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-666

Query Match
Best Local Similarity 7.1%; Score 36.4; DB 15; Length 6222;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 344 AATATCAATATCAGTGGCAGCTAACTGCTCCAGACCTTACACTTAAAGCAATCTCTCTTA 403
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5664 AATTATACCTAATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5605
QY 404 CTCACCTTCTCCAGACAGCAAGCAAAATACAAAACAGATGCCAAAATCTCCAACTTA 463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5604 CTTTATATCTTTTCTACTATAATAATCAACAATACATACATATATACTAAACCCCTTA 5545
QY 464 CTGCGACTACCCAAAATTTCCACTGATATT 493
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5544 TAAATCCTTACTAAAAAATCAACAATCTAT 5515
```

```
RESULT 14
US-10-027-632-135035
; Sequence 135035, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135035
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-135035
```

```
Query Match
Best Local Similarity 7.1%; Score 36.2; DB 13; Length 777;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 371 CTCAGACCTTACACCTTAAAGCAATCTCTTACTTCACTTCCAGAACAGCAA 430
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
322 CTCAGCCTGGGCAAGTCAGACGACTTCCATCTAAACACACCACCACCAACATAAA 381
QY 431 ATACAAAAACAGATGCCAAAATCTCCACACTACTGC 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 AAACAAAAACAAACAAAAAAACCTTCCCTCTTTC 418
```

```
RESULT 15
US-10-027-632-135036
; Sequence 135036, Application US/10027632
; Publication No. US20020198371A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135036
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-135036
```

```
Query Match
Best Local Similarity 7.1%; Score 36.2; DB 13; Length 777;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 371 CTCAGACCTTACACCTTAAAGCAATCTCTTACTTCACTTCCAGAACAGCAA 430
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
322 CTCAGCCTGGGCAAGTCAGACGACTTCCATCTAAACACACCACCACCAACATAAA 381
QY 431 ATACAAAAACAGATGCCAAAATCTCCACACTACTGC 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 AAACAAAAACAAACAAAAAAACCTTCCCTCTTTC 418
```

Search completed: July 30, 2004, 20:26:37
Job time : 413 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 17:06:36 ; Search time 2908 Seconds

(without alignments)
5237.177 Million cell updates/sec

Title: US-09-386-850-7_COPY_1_510

Perfect score: 510

Sequence: 1 AAACGACTGACGAGAGC.....TTTTCAGGTCGGTGCATA 510

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pig:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472.8	92.7	483	10	BF562410
2	399.8	78.4	562	10	AW533278
3	390.8	76.6	402	14	CB769246
4	152.8	30.0	634	10	BF534987

ALIGNMENTS

RESULT 1	BF562410	483 bp	linear	EST 12-DEC-2000
LOCUS	UI-R-BUO-anf-a-06-0-UI.r1	UI-R-BUO	Rattus norvegicus	cDNA clone
DEFINITION	UI-R-BUO-anf-a-06-0-UI 5', mRNA sequence.			
ACCESSION	BF562410			
VERSION	BF562410.1	GI:11672140		
KEYWORDS	EST.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 483)			
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res. 6 (9), 791-806 (1996)			
MEDLINE	97044477			
PUBMED	889548			
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu			

5	152.2	29.8	697	12	BG872108	BG872108	602792936
6	151.2	29.6	581	10	BF539002	BF539002	602048549
7	151.2	29.6	605	10	BF540223	BF540223	602050271
8	151.2	29.6	644	10	BF537882	BF537882	602049321
9	151.2	29.6	652	12	BG865810	BG865810	602798210
10	151.2	29.6	676	10	BF299863	BF299863	602030493
11	149.6	29.3	872	10	BF534149	BF534149	602047633
12	149.2	29.3	652	10	BF302989	BF302989	602030914
13	148	29.0	645	10	BF540321	BF540321	602052377
14	148	29.0	655	10	BF540217	BF540217	602050262
15	147.2	28.9	653	10	BF536022	BF536022	602054151
16	146.4	28.7	481	10	BF534360	BF534360	602047860
17	146.4	28.7	513	12	BG872364	BG872364	6020792651
18	144.8	28.4	745	10	BF302398	BF302398	602031395
19	144	28.2	515	10	BF299995	BF299995	602030654
20	144	28.2	649	10	BF535583	BF535583	602051874
21	144	28.2	649	10	BF538163	BF538163	602053634
22	144	28.2	670	10	BF535602	BF535602	602051893
23	144	28.2	674	10	BF537472	BF537472	602053115
24	144	28.2	867	10	BF540227	BF540227	602050275
25	143.6	28.2	624	10	BF300221	BF300221	602032126
26	143.6	28.2	644	12	BG865669	BG865669	602783834
27	143.2	28.1	728	10	BF300831	BF300831	602028829
28	143.2	28.1	942	12	BG865578	BG865578	602785548
29	143	28.0	660	10	BF537022	BF537022	602048926
30	142.6	28.0	942	10	BF302832	BF302832	602032723
31	142.4	27.9	553	12	BG869030	BG869030	602784661
32	142.4	27.9	685	10	BF535415	BF535415	602051809
33	142	27.8	645	10	BF300735	BF300735	602031918
34	142	27.8	648	10	BF536686	BF536686	602048662
35	142	27.8	648	12	BG872654	BG872654	602793712
36	142	27.8	650	12	BG870684	BG870684	602791670
37	142	27.8	654	10	BF302080	BF302080	602033029
38	142	27.8	655	10	BF537445	BF537445	602050084
39	142	27.8	660	12	BG872510	BG872510	602793545
40	142	27.8	661	11	BQ035428	BQ035428	Mus muscu
41	142	27.8	663	10	BF539051	BF539051	602051406
42	142	27.8	664	10	BF538737	BF538737	602051075
43	142	27.8	668	12	BG865577	BG865577	602783725
44	142	27.8	690	10	BF302310	BF302310	602031281
45	142	27.8	695	10	BF540179	BF540179	602050223

cdNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LMML (info@lmml.gov). IMAGE ID= 1801619
Seq primer: M13 Forward.

FEATURES

Location/Qualifiers

```
1. 483
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BU0-anf-a-06-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-BU0"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BU0
library is a subtracted library derived from a mixture of
eye and ganglia tissues. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratest.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)"
```

ORIGIN

Query Match 92.7%; Score 472.8; DB 10; Length 483;
Best Local Similarity 99.6%; Pred. No. 1.1e-111; Indels 0; Gaps 0;
Matches 474; Conservative 0; Mismatches 2;

```
QY 23 CTGACCAGCACATTTCCCGCTCAGAGTTTCTCCAAAGGGGTACCAAGATCAAGTCAC 82
DB 8 CTGACCAGCACATTTCCCGCTCAGAGTTTCTCCAAAGGGGTACCAAGATCAAGTCAC 67
QY 83 TGTATTGATCTTTGGCTGTGGATCTCTTAGCATGCTTCCAGTCAGTGAGGTGTCA 142
DB 68 TGTATTGATCTTTGGCTGTGGATCTCTTAGCATGCTTCCAGTCAGTGAGGTGTCA 127
QY 143 GAGGCCCAAGAGACACATATCTTAGAAGACCAACAGATCTTCAACTCTTCTCTCAT 202
DB 128 GAGGCCCAAGAGACACATATCTTAGAAGACCAACAGATCTTCAACTCTTCTCTCAT 187
QY 203 ATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATAGGAGTAAACATCACTATACCT 262
DB 188 ATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATAGGAGTAAACATCACTATACCT 247
QY 263 TAAATCTTCAACCACTCTGTCTTGTAAATCTTCCCGTTTATCACTGGACCACTAT 322
DB 248 TAAATCTTCAACCACTCTGTCTTGTAAATCTTCCCGTTTATCACTGGACCACTAT 307
QY 323 TGTGTGTACAAGTACCACTGAATATCAATATCAGTGGCAGCTAACTGCTCCAGACCTTA 382
DB 308 TGTGTGTACAAGTACCACTGAATATCAATATCAGTGGCAGCTAACTGCTCCAGACCTTA 367
QY 383 CACCTCTAAGCAATCTCTTACTCAACTTCATTCACAGAACAGCAAGCAATACAAAACAG 442
DB 368 CACCTCTAAGCAATCTCTTACTCAACTTCTTCCACAGAACAGCAAGCAATACAAAACAG 427
QY 443 ATGCCAAATCTCCACACTCTGACACTACCCAAATCCCTGATATTTTGA 498
DB 428 ATGCCAAATCTCCACACTCTGACACTACCCAAATCCCTGATATTTTGA 483
```

RESULT 2

AW533278/c
LOCUS
DEFINITION
UI-R-BU0-anf-a-06-0-UI.s1 UI-R-BU0 Rattus norvegicus cDNA clone
UI-R-BU0-anf-a-06-0-UI 3', mRNA sequence.
AW533278
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS
TITLE

JOURNAL

MEDLINE

PubMed

COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 562)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized ganglia library cDNA library Preparation: M.B. Soares
Lab Clone distribution: clones will be available through Research
Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

Location/Qualifiers

```
1. 562
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BU0-anf-a-06-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-BU0"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BU0
library is a subtracted library derived from a mixture of
eye and ganglia tissues. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratest.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_TISSUE=ganglia
TAG_LIB=UI-R-BU0
TAG_SEQ=GCAGAC
```

ORIGIN

Query Match 78.4%; Score 399.8; DB 10; Length 562;
Best Local Similarity 99.5%; Pred. No. 8.9e-93;
Matches 401; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 108 CCTTCTAGCATCTTCCAGTCAGGTGAGGTCTCAGAGCCCAAGACACATAATCC 167
DB 562 CCTTCTAGCATCTTCCAGTCAGGTGAGGTCTCAGAGCCCAAGACACATAATCC 503
QY 168 TAGAGACAAACAGATCTTCAACTCTTCTCATTTATCTTGTCTTTCAGCCTGATCCAA 227
DB 502 TAGAGCCCAACAGATCTTCAACTCTTCTCATTTATCTTGTCTTTCAGCCTGATCCAA 443
QY 228 TGTGTGACAAATAGGAGTAAACATCACTATACCTTAAATCTTCAACCACTCGTGTCT 287
DB 442 TGTGTGACAAATAGGAGTAAACATCACTATACCTTAAATCTTCAACCACTCGTGTCT 383
QY 288 TGTTAATCTTCCCGTTTATCACTGGACCACTTGTGTACAGAGTACCCTGAATA 347
DB 382 TGTTAATCTTCCCGTTTATCACTGGACCACTTGTGTACAGAGTACCCTGAATA 323
QY 348 TCAATATCAGTGGCAGCTAACTGCTCCAGACCCCTACACCTTAAAGCAATCTCTCTACTCA 407
DB 322 TCAATATCAGTGGCAGCTAACTGCTCCAGACCCCTACACCTTAAAGCAATCTCTCTACTCA 263
```


QY 408 ACTTATTCCACAGAACCAATACAAAACAGATGCCAAATCTCCAACTACTGC 467
|||||
Db 262 ACTTCTTCCACAGAACCAATACAAAACAGATGCCAAATCTCCAACTACTGC 203
|||||
QY 468 GACTACCCAAAATTCCTGATATTTTGAAGGTGGTGGCAA 510
|||||
Db 202 GACTACCCAAAATTCCTGATATTTTGAAGGTGGTGGCAA 160
|||||
RESULT 3
CB769246 402 bp mRNA linear EST 16-MAY-2003
LOCUS AMGNNUC:SRPB2-00151-E10-A.srbp2 (10220) Rattus norvegicus cDNA
DEFINITION clone srbp2-00151-e10 5', mRNA sequence.
ACCESSION CB769246
VERSION CB769246.1 GI:29857637
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
AUTHORS Rattus norvegicus
TITLE Rattus norvegicus
JOURNAL Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00151 row: e column: 10.

FEATURES
Location/Qualifiers
1..402
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srbp2-00151-e10"
/tissue_type="prostate tissue"
/clone_lib="srbp2 (10220)"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat prostate normalized double selected poly(A+) mRNA size fraction > 1 kb"

ORIGIN
Query Match 76.6%; Score 390.8; DB 14; Length 402;
Best Local Similarity 99.5%; Pred. No. 1.8e-90;
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 CTGACTGACGAGAGCTTCTGACACGACATTTCCCGCTCAGAAAGTTTCTCCAGGGG 63
Db 9 CCGACTGACGAGAGCTTCTGACACGACATTTCCCGCTCAGAAAGTTTCTCCAGGGG 68
QY 64 CTACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCTCTAGCATCTTC 123
Db 69 CTACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCTCTAGCATCTTC 128
QY 124 CAGTCAGGTGAGGTGTCAGAGCCCAAGAACACATATCTTAGAGACAAAGAT 183
Db 129 CAGTCAGGTGAGGTGTCAGAGCCCAAGAACACATATCTTAGAGACAAAGAT 188
QY 184 CTTTCAACTCTTCTCATTTATCTTTGCTTTCAGCTGATCCCAATGCTGACAAATAGGA 243
Db 189 CTTTCAACTCTTCTCATTTATCTTTGCTTTCAGCTGATCCCAATGCTGACAAATAGGA 248
QY 244 GTAACATCACTATACCTTAAATCTTCAACCACTCGTGTCTTCTTAAATCTTCCCGGT 303
Db 249 GTAACATCACTATACCTTAAATCTTCAACCACTCGTGTCTTCTTAAATCTTCCCGGT 308
QY 304 TTTATCACTGGACCACTGTTGTACAGGTACCACTGAATATCAATATCAGTGGCAG 363
Db 309 TTTATCACTGGACCACTGTTGTACAGGTACCACTGAATATCAATATCAGTGGCAG 368
QY 364 CTAAGTCTCCAGACCTTACCTCTTAAGCAATC 397

Db 369 CTAAGTCTCCAGACCTTACACCTTAAAGCAATC 402
|||||

RESULT 4
BF534987 634 bp mRNA linear EST 11-DEC-2000
LOCUS 602050440F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4190048 5',
DEFINITION mRNA sequence.
ACCESSION BF534987
VERSION BF534987.1 GI:11622350
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 634)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9516 row: o column: 09
High quality sequence stop: 614.

FEATURES
Location/Qualifiers
1..634
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4190048"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 30.0%; Score 152.8; DB 10; Length 634;
Best Local Similarity 62.4%; Pred. No. 8.3e-29;
Matches 318; Conservative 0; Mismatches 167; Indels 25; Gaps 4;
QY 1 AAACCTGACTGACGAGAGCTTCTGACACGACATTTCCCGCTCAGAAAGTTTCTCCAAG 60
Db 5 AAAATAACTGACGAGAGCTTCTGACACGACATTTCCCGCTCAGAAAGTTTCTCCAAG 64
QY 61 GGGCT-ACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCTCTAGCATG 119
Db 65 AGGCTGAAAAAAGATGAAGCCCAATTTGCTTTAGGCTCTGCTTTGTTAGGATG 124
QY 120 CTTTCAGTCAAGTGAAGGTGTCAGAGGCCCAAGAACACATATCTTAGAGACAA 179
Db 125 CTTTCGTGCTGTGAGTGTACAGAGGCCCTTAGAAG---ACATGATCTTAGAGACCAAT 181
QY 180 AGATCTTTCAATCTTCTCTCATTTATCTTTGGCTTTCAGCTGATCCCAATGTTGGACAAAT 239
Db 182 CCT 226
QY 240 AGGATTAACATCACTATACCTTAAATCTTCAACCACTCGTGTCTTCTTAAATCTTTC 299
Db 227 CCTCCACCTTTTGGTCCAGGAATTTGGTAGACCACTCCCTTCCAGGAT 286
QY 300 CGGTTTTTATCTGACCACTGTTGTTGTACAGGTACCACTGAATATCAATATCAGT 359
Db 287 TGGTCGAAGCGCCCTTCCACCTCTCTTGTCCACCACTTCTCTCCACATCTTAGACCTCCAAG 346

```
QY 360 GCAGTAAGTCTCCAGACCCCTACACCTCTAAGCAATCCTCTACTCACTTCACTCCAC 419
Db 347 CAATCGGTCTCTCTCA-----CCTACTCCAGCAATCTCTCTACCGGACCTCTTACCAC 400
QY 420 AGAACAGCAATACAAAAACAGATGCCAAATCTCCAAACACTACTGCGCACTTACCCAAA 479
Db 401 AGTACAAGCACTACAAATGCCAGCTGCCAGATCTCCATACTACTCTCTACTGCTAGAGA 460
QY 480 TTCCACTGATATTTTGAAGGTGGTGGCAA 509
Db 461 TTCCACTGATATTTTGGAGGTGGTGGGAA 490

RESULT 5
BG872108 697 bp mRNA linear EST 29-MAY-2001
LOCUS 602792936F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4924089 5',
DEFINITION mRNA sequence.
ACCESSION BG872108
VERSION BG872108.1 GI:14222648
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 697)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10846 row: h column: 10
High quality sequence stop: 635.
Location/Qualifiers
1..697
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4924089"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Not1; Site_2: Sal1; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
Location/Qualifiers
1..697
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4924089"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Not1; Site_2: Sal1; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 29.8%; Score 152.2; DB 12; Length 697;
Best Local Similarity 62.3%; Pred. No. 1.2e-28;
Matches 317; Conservative 0; Mismatches 168; Indels 24; Gaps 4;

QY 2 AACTGACTGACCAAGAGCTTGTGACAGACATTTCCCGCTCAGAAGTTTCTCCAAG 61
Db 1 AAATACTGACCAAGCGCTTGTGACAGACATTCACAGCTAAAGAGTGTCTCCAAGA 60

QY 62 GGCT-ACCAAGATGAAGTCACTGTATTGATCTTTGGCTGTGGATCTCTAGCATGC 120
Db 61 GGCTGAAAAGATGAAGCCACTCAATTTGGTCTTAGGCTCTGTCATTTGTAGGATGC 120

QY 121 TTCAGTCAAGTGAAGGTGTGACAGGCCCAAGAGCAACATAATCTTAGAGCAACAA 180
Db 121 TTCTGTGCTGTGAGTGTGACAGGCCCTAGAG---ACATGATCTAGAGCCATTC 177

QY 181 GATCCTTCAACTCTTCTCTCATTTATCTTGGTCTTTCAGCCTGATCCCAATGTGCAATA 240
```

```
Db 178 CCTCTCTCTCCCTCCCATGGTCCAGGAATT---GGTAGACCACACCCCTCCACCTTT 234
QY 241 GGAGTAACAATCACTATACCCCTTAAATCTTCAACACACCTGTTCTTGTAACTTCCC 300
Db 235 GGTCCAGGAATTTGGTAGACACCCCTCCACCCCTTTGGTCCAGGAATTTGGTCCGACCCC 294
QY 301 GGTTTTATCACTGGACCAACCATTTGGTTGTACAAGGTACCACTGAATATCAATATCAGTGG 360
Db 295 CCTCMACTCTTGTCCACAGGTTCTCTCCATCTAGACCT-----C 337
QY 361 CAGCTAACTGCTCCAGACCCCTACACCTCTAAGCAATCTCTCTACTCAACTTCACTCCACA 420
Db 338 CAAGCAATCTCGTCTCCATCCACTACTCCAAGCATCTCTCTACCGACCTCTTACCACA 397
QY 421 GAACAGCAATATACAAAAACAGATGCCAAATCTCCAAACACTACTGCGCACTTACCCAAA 480
Db 398 GTACAAGCAACTACAAATGCCAGTCCAGTATCTCCATACTACTCTCTACTCTAGAT 457
QY 481 TCCACTGATATTTTGAAGGTGGTGGCAA 509
Db 458 TCCACTGATATTTTGGAGGTGTGGAA 486

RESULT 6
BF539002 581 bp mRNA linear EST 11-DEC-2000
LOCUS 602048549F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4187797 5',
DEFINITION mRNA sequence.
ACCESSION BF539002
VERSION BF539002.1 GI:11626383
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 581)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9511 row: a column: 14
High quality sequence stop: 580.
Location/Qualifiers
1..581
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4187797"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
Location/Qualifiers
1..581
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4187797"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 29.8%; Score 151.2; DB 10; Length 581;
Best Local Similarity 62.2%; Pred. No. 2.1e-28;
Matches 317; Conservative 0; Mismatches 168; Indels 25; Gaps 4;

QY 1 AAATCACTGACCAAGAGCTTCTGACAGCAATTTCCCGCTCAGAAGTTTCTCCAAG 60
Db 31 AAATACTGACCAAGCCGCTTCTGACAGCAATTCACAGCTAGAGATGTCTCAAG 90
```



```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="TMAG:4188749"
/lab_host="DH10B (T1 phage)"
/clone_lib="MCT CGAP SC2"

```

ORIGIN

Query Match	29.6%;	Score	151.2;	DB	10;	Length	644;
Best Local Similarity	62.2%;	Ident.	No. 2.2e-28;				
Matches	317;	Conservative	0;	Mismatches	168;	Indels	25; Gaps 4;

1	AAACTGACTGACACAGAGCTTCTGACACAGACATTTCCCGCTCAGAAAGTTCTTCCAAG	60
yy		
bb		
3	AAATAACTGACACAGACCGTTCTGACACAGACATTCACAGCTAAGAAGTGTCTCCAAG	62
yy		
bb		
61	GGGCT-ACCAAGATGAAGTCACTGTAATTGATCTTTGGCGCTGGATCCCTTCTAGCATG	119
yy		
bb		
63	AGGCTGAAAAAGATGAAGCCACTCAATTGGTCTTAGGCTCTGCATTTCTTGTAGGATG	122
yy		
bb		
120	CTTCCAGTCAGGTGAGGTGTGACAGGCCACAGAGACACATATCTCTAGNAGACACA	179
yy		
bb		
123	CTTCCCTGTCTGTGAGTGTACAGAGAGGCCCTAGAAG--ACATGATCTCTAGAGACCAAT	179
yy		
bb		
180	AGATCCTTCAACTCTTCCCTCATTTCTTGGTCTTCAGCCTGATCCCAATGGTGGACAAT	239
yy		
bb		
180	CCCTCTCTCCGCCCTCCCATGGTCCAGAAAT---GGTAGACACACACCCCTCCACCCCTT	236
yy		
bb		
240	AGGAGTAACAATCACTATACCTTAAATCTTCAACACCTCGTGTCTTGTGTTAAATCTTCC	299
yy		
bb		
237	TGTTCCAGGAAATGGTAGACACACCCCTCCACCCCTTGGTCCAGGAAATGGTCGACCACC	296
yy		
bb		
300	CGGTTTTATCACTGGAACCAATGTTGGTGTGAACAGGTACCACTGAATATCAATATCATGTG	359
yy		
bb		
297	CCCTCCACCTCTCTGTGCCACAGTTCCTGCCATCTCTAGACCT-----	339
yy		
bb		
360	GCAGCTTAAGTGTCCAGACCCCTACACCTCTAGCAATCTCTACTCAACTTCATTCCAC	419
yy		
bb		
340	-CCAAGCAATCCGCTCTCCTCCACTACTCCAGCATTCCTCTACGGACCTCTCTACAC	398
yy		
bb		
420	AGAACAGCAAAATACAAAAACAGATGCAAAAATCTCCAAACACTACTCGCACTACCCAAA	479
yy		
bb		
399	AGTACAGCAACTACATGCCACTGCCAGTATCTCCATTAATCTCTACTCTGCTAGAGA	458
yy		
bb		
480	TTCCACTGATATTTTTGAAGGTCGTGGCAA	509
yy		
bb		
459	TTCCACTGATATTTTTGAAGTGTGGAA	488
yy		
bb		

RESULT 9	
BG865810	
LOCUS	652 bp mRNA linear EST 29-MAY-2001
DEFINITION	602788210F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4914002 5', mRNA sequence.
ACCESSION	BG865810
VERSION	BG865810
KEYWORDS	GI:14216350
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 652)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapsb-r@mail.nih.gov Tissue Procurement: Jeffery E. Green, M.D.

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10820 row: d column: 03
High quality sequence stop: 636.
Location/Qualifiers
1. 652
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4914002"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_S62"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
Not1; Site_2: Sal1; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

ORIGIN

Query Match	29.6%;	Score 151.2;	DB 12;	Length 652;
Best Local Similarity	62.2%;	Pred. No. 2.2e-28;		
Matches 317;	Conservative	0;	Mismatches 168;	Indels 25; Gaps 4;

Qy	1	AAACTGACTCACCAGAGAGCGTCTCTGACACGACACATTTCCCGCTCAGAAAGTTTCTCCAAG	60
Db	3	AAATTAATGACACAGACCGCTTCTGACACGACACATTCACACGATGAAGAGTGCTCCAAG	62
Qy	61	GGGCT-ACCAAGATGAAGTCACTGTATTGTATCTTTGGCTGTGGATCCTTCTAGCATG	119
Db	63	AGGCTGAAAAAAGATGAAGCACTCAATTTGGTTAGGCTCTGCAATCTTTGTAGGATG	122
Qy	120	CTTCCAGTCAGGTGAGGGTGTTCAGAGGCCCAAGAACACATAATCCTAGACACAACA	179
Db	123	CTTCCCTGTCTGTGAGTGTACAGAGGCCCTAGAG--ACATGATCCTAGAGACCATT	179
Qy	180	AGATCCTTTCAACTCTTCCCTCATTTATCTTGTCTTTCAGCCTGATCCCAATGGTGGACAAAT	239
Db	180	CCCTCCTCTCCCTCCCTCCCATGTCTCAGGAATT---GGTAGACACACACCTCCACCGTT	236
Qy	240	AGGAGTAACATACTACTATACGCTTAATCTTCAACACCTCGTGTCTTGTGTAATCTTCC	299
Db	237	TGCTCCAGGAATTGGTAGACACCCCTCCACCCCTTTGGTCCAGGAATTTGTTCGACACCC	296
Qy	300	CGGTPTTTATCACTGGACACCAATCGTGTGTACAAGGTACCACTCAATATCAATATCATGT	359
Db	297	CCCTCCACCTCCTTGTCCACCGAGTCTCTCCACATCCTAGACCT-----	339
Qy	360	GCAGCTAACTGTCTCCAGACCTCAACCTCTAAGCAATCTCTCTACTCAACTTCAATCCAC	419
Db	340	-CCAAGCAATCCGTCTCTCTCCAGTACTTCCAAGCATTTCTCTACCGGACCTCTTACAC	398
Qy	420	AGACACAGCAATACAAAACAGATGCCAAATCTCCAACACTACTGCGACTACCCAAA	479
Db	399	AGTACAAGCAACTCAATGCCAGTCCGAGTATCTCCATACTACTCTCTACTGCTAGGA	458
Qy	480	TTCCACTGATATTTTGAAGGTGGTGGCAA	509
Db	459	TTCCACTGATATTTTGGAGTTTGGGAA	488

RESULT 10	BF299863	676 bp	mRNA	linear	EST 21-NOV-2000
LOCUS	602030493P1	NCI_CGAP_SG2	Mus musculus	cDNA	clone IMAGE:4165726 5',
DEFINITION	mRNA sequence.				
ACCESSION	BF299863				
VERSION	BF299863.1	GI:11246386			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 676)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM9453 row: i column: 23
High quality sequence stop: 626.
Location/Qualifiers
1..676
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4165726"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

FEATURES
source
1..676
Query Match 29.6%; Score 151.2; DB 10; Length 676;
Best Local Similarity 62.2%; Pred. No. 2.2e-28;
Matches 317; Conservative 0; Mismatches 168; Indels 25; Gaps 4;
QY 1 AAATGACTGACGACGAGAGCTTCTGACGACGACATTTCCCGCTCAGAGTTTCTTCCAAG 60
Db 1 AAATAACTGACGACGACGCGTCTTGACGACGACATTCACGAGTAAGAGTGTCTCCAAG 60
QY 61 GGCGCT-ACCAAAGATGAAGTCACTGTATTTGATCTTTGGCTGTGGATCCTTCTAGCATG 119
Db 61 AGGCTGAAAAGATGAAGCCACTCAATTGGTCTTAGGCTCTGCATCTTGTAGATG 120
QY 120 CTTCCAGTCAGGTGAGGGGTGTGAGAGCCCAAGAGCAACATAATCTTAGAGAGCAACA 179
Db 121 CTTCCCTGCTGTGTGAGTGTACAGAGGCCCTAGAG---ACATGATCCTAGAGGACCAT 177
QY 180 AGATCCTTCAACTCTTCTCATTTATCTGTCTTTCAGCCTGATCCCAATGGTGGACAAAT 239
Db 178 CCCTCCTCTCCCTCCCGATGTCAGGAAT---GGTAGACACACCCCTCCACCTT 234
QY 240 AGAGTAAACAATCACTATACCTTAAATCTTCAACACCTCTGTGTCTTGTAACTTCC 299
Db 235 TGGTCCAGGAATGTGAGACACCCCTCCACCTTTGGTCAGGAATGGTTCGACCCACC 294
QY 300 CGGTTTTATCACTGACGACCACTTGGTGTACAGGTACCACTGCACTACCCACAAA 359
Db 295 CCCTCCACTCTCTGTCCACCACTTCTCCACATCCTAGACCT----- 337
QY 360 GCAGCTAACTGCTCCAGACCCCTACCTCTTAAGCAATCTCTACTCACTTCACTCCAC 419
Db 338 -CCAAGCAATCCGTCCTCCACTACTCAAGCAATCTCTCCGACCCCTCTTACCAC 396
QY 420 AGAACAAGCAATACAAAACAGATGCCAAATCTCCAACTACTGCGACTACCCAAA 479
Db 397 AGTACAAGCAACTACAATGCCAGTCCAGTATCTCCATACTACTCTACTGCTAGAG 456
QY 480 TTCCACTGATATTTTGAAGGTGGGCAA 509
Db 457 TTCCACTGATATTTTGGAGGTGTGGGAA 486

RESULT 11
LOCUS BF534149
DEFINITION 602047633F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4187024 5',
mRNA sequence.
ACCESSION BF534149
VERSION BF534149.1 GI:11621512
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 872)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM9509 row: a column: 09
High quality sequence stop: 621.
Location/Qualifiers
1..872
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4187024"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 29.3%; Score 149.6; DB 10; Length 872;
Best Local Similarity 62.0%; Pred. No. 5.9e-28;
Matches 316; Conservative 0; Mismatches 169; Indels 25; Gaps 4;
QY 1 AAATGACTGACGACGAGAGCTTCTGACGACGACATTTCCCGCTCAGAGTTTCTTCCAAG 60
Db 1 AAATAACTGACGACGACGCGTCTTGACGACGACATTCACGAGTAAGAGTGTCTCCAAG 60
QY 61 GGCGCT-ACCAAAGATGAAGTCACTGTATTTGATCTTTGGCTGTGGATCCTTCTAGCATG 119
Db 61 AGGCTGAAAAGATGAAGCCACTCAATTGGTCTTAGGCTCTGCATCTTGTAGATG 120
QY 120 CTTCCAGTCAGGTGAGGGGTGTGAGAGCCCAAGAGCAACATAATCTTAGAGAGCAACA 179
Db 121 CTTCCCTGCTGTGTGAGTGTACAGAGGCCCTAGAG---ACATGATCCTAGAGGACCAT 177
QY 180 AGATCCTTCAACTCTTCTCATTTATCTGTCTTTCAGCCTGATCCCAATGGTGGACAAAT 239
Db 178 CCCTCCTCTCCCTCCCGATGTCAGGAAT---GGTAGACACACCCCTCCACCTT 234
QY 240 AGAGTAAACAATCACTATACCTTAAATCTTCAACACCTCTGTGTCTTGTAACTTCC 299
Db 235 TGGTCCAGGAATGTGAGACACCCCTCCACCTTTGGTCAGGAATGGTTCGACCCACC 294
QY 300 CGGTTTTATCACTGACGACCACTTGGTGTACAGGTACCACTGCACTACCCACAAA 359
Db 295 CCCTCCACTCTCTGTCCACCACTTCTCCACATCCTAGACCT----- 337
QY 360 GCAGCTAACTGCTCCAGACCCCTACCTCTTAAGCAATCTCTACTCACTTCACTCCAC 419

```

Db      338 -CCAGCAATCGTCTCTCCACCTACTCCAGCAATTCCTCTACCGGAGCTCTCCATCCAC 396
Qy      420 AGAACAAGCAATAAACAACAGATGCCAAATCTCCACACTACTGCGACTACCCAAA 479
Db      397 AGTACAAGCAACTACAATGCCAGCTGCCAGTATCTCCATACTACTCTACTGCTAGAGA 456
Qy      480 TTCCACTGATATTTTGAAGGTGGGCAA 509
Db      457 TTCCACTGATATTTTGGAGGTGGGAA 486

RESULT 12
BF302989
LOCUS   BF302989
DEFINITION BF302989
          652 bp mRNA linear EST 21-NOV-2000
          602030914F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4166139 5',
          mRNA sequence.
ACCESSION BF302989
VERSION   BF302989
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 652)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM9454 row: k column: 04
          High quality sequence stop: 645.
          Location/Qualifiers
FEATURES
source   1..652
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="FVB/N"
          /db_xref="taxon:10090"
          /clone="IMAGE:4166139"
          /lab_host="DH10B (T1 phage-resistant)"
          /clone_lib="NCI CGAP_SG2"
          /clone_host="NCI CGAP_SG2"
          /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
          NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
          dt. Average insert size 1.3 kb. Constructed by Life
          Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 29.3%; Score 149.2; DB 10; Length 652;
Best Local Similarity 61.8%; Pred. No. 7.2e-28;
Matches 315; Conservative 0; Mismatches 170; Indels 25; Gaps 4;

Qy      1 AAATCACTGACACAGAGAGCTTCTGACACGACATTTCCCGCTCAGAGTCTTCCAAAG 60
Db      5 AAATAAAGTACACAGACCGCTTCTGACACGACATTCACGAGTAAAGAGTCTTCCAAAG 64
Qy      61 GGCT-ACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCCTCTAGCATG 119
Db      65 AGCTGAAAGAGATGAGCCACTCAATTTGGCTTAGGCTCTGCATTTCTGTAGATG 124
Qy      120 CTTCCAGTCAAGGTGTGAGAGGCCCAAGACACATATCTTCTAGAGACACACA 179
Db      125 CTTCCCTGCTGAGTGTGACAGAGGCCCTAGAG---ACATGATCTCTAGAGACCAT 181
Qy      180 AGATCCTTCAACTCTTCTCATTAATCTTGCTTTCAGCCCTGATCCCAATGGTGACAAAT 239
Db      182 CCCTCTCTCTCCCTCTCCCATGTTCAGGAANT--GGTAGACACACCTCCACCTT 238

```

```

Qy      240 AGGAGTAACAATCACTATACCTTAAATCTTCAACCACTCGTGTGTGTGTTAAATCTTCC 299
Db      239 TGGTCCAGGAATGGTAGACACCCCTCCACCTGTGGTCCAGGAATTGGTCCGACACC 298
Qy      300 CGGTTTTATCACTGACACCACTTGGTTGTACAGGTACCACTGAATATCAATATCAGTG 359
Db      299 CCTCCACCTCTGTGTCACCACTGTTCTCCACATCTCTAGACCT----- 341
Qy      360 CGAGCTAACTGCTCCAGACCTTACACCTCTTAAGCAATCTCTACTCACTCACTTCAATCCAC 419
Db      342 -CCAAGCAATCGTCTCTCTCCACTACTCCAGCAATTCCTCTACCGGACCTCTCTACAC 400
Qy      420 AGAACAAGCAATAAACAACAGATGCCAAATCTCCAACTACTGCGACTACCCAAA 479
Db      401 AGTACAAGCAACTACAATGCCAGCTGCCAGTATCTCCATACTACTCTCTACTGCTAGAGA 460
Qy      480 TTCCACTGATATTTTGAAGGTGGGCAA 509
Db      461 TTCCACTGATATTTTGGAGGTAGTGGAA 490

RESULT 13
BF540321
LOCUS   BF540321
DEFINITION BF540321
          645 bp mRNA linear EST 11-DEC-2000
          602052377F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4191474 5',
          mRNA sequence.
ACCESSION BF540321
VERSION   BF540321
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 645)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM9520 row: j column: 19
          High quality sequence stop: 638.
          Location/Qualifiers
FEATURES
source   1..645
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="FVB/N"
          /db_xref="taxon:10090"
          /clone="IMAGE:4191474"
          /lab_host="DH10B (T1 phage-resistant)"
          /clone_lib="NCI CGAP_SG2"
          /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
          NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
          dt. Average insert size 1.3 kb. Constructed by Life
          Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 29.0%; Score 148; DB 10; Length 645;
Best Local Similarity 61.8%; Pred. No. 1.5e-27;
Matches 315; Conservative 0; Mismatches 170; Indels 25; Gaps 4;

Qy      1 AAATCACTGACACAGAGAGCTTCTGACACGACATTTCCCGCTCAGAGTCTTCCAAAG 60
Db      1 AAATAAAGTACACAGACCGCTTCTGACACGACATTCACGAGTAAAGAGTCTTCCAAAG 60
Qy      61 GGCT-ACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCCTCTAGCATG 119

```

Db 61 AGGCTGAAAAAAGATGAAGCACTCAATTTGGTCTTAGGCTCTGCATCTTTGTAGGATG 120
QY 120 CTTCCAGTCAGGTGAGGTGTACAGGCCCAAGAGCAACATAATCTCTAGAACAAACA 179
Db 121 CTTCCAGTCAGGTGAGGTGTACAGGCCCAAGAGCAACATAATCTCTAGAACAAACA 177
QY 180 AGATCCCTTCAACTCTCTCTCATATCTTTGGTCTTAGGCTCTGCATCTTTGTAGGATG 239
Db 178 CCCTCCCT 234
QY 240 AGGAGTAACTACTATACCTTTAAATCTTTCAACCACTCTCTCTCTCTCTCTCTCTCTCT 299
Db 235 TGGTCCAGGAAATTTGAGAGCAACCCCTCCACCCCTTTGGTCCAGGAATTTGGTCCAGCACC 294
QY 300 CGGTTTTATCACTGGACCACTATTTGGTGTACAGGTACCACTGAATATCAATATCAGTG 359
Db 295 CCCTCCACCT 337
QY 360 GCAGCTAACTGCTCCAGACCCCTACACCTCTAAGCAATCTCTCTCTCTCTCTCTCTCTCTCT 419
Db 338 -CCAAGCAATCCGCT 396
QY 420 AGAACAGCAAAATACAAAACAGATGCCAAATCTTCAACACTACTCGGACTACCCAAAA 479
Db 397 AGTACAGCAACTACTACTGCGCAGTGCAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 456
QY 480 TTCCACTGATATTTTGAAGGTGGTGCAC 509
Db 457 TTCCACTGATATTTTGAAGGTGGTGGAA 486

RESULT 14
BF540217 655 bp mRNA linear EST 11-DEC-2000
LOCUS 602050262F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4189599 5',
DEFINITION mRNA sequence.
ACCESSION BF540217
VERSION BF540217.1 GI:11627598
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 655)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9515 row: 1 column: 16
High quality sequence stop: 637.
Location/Qualifiers
1. .655
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4189599"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

Query Match 29.0%; Score 148; DB 10; Length 655;
Best Local Similarity 61.8%; Pred. No. 1.5e-27;
Matches 315; Conservative 0; Mismatches 170; Indels 25; Gaps 4;
QY 1 AACTGCTGCTACACAGAGGCTTCTGACAGCAACATTTCCCGCTCAGAAAGTTTCTCCAAG 60
Db 5 AATAAATACTACACAGACCGCTTCTGACAGCAACATTTCCAGAGTGTCTCCAAG 64
QY 61 GGGCT-ACCAAAAGATGAAGTCACTGATTTTGAATCTTTGGCTGTGGATCTTCTAGCATG 119
Db 65 AGGCTGAAATAAGATGAGCCACTCAATTTGGTCTTAGGCTCTGCATCTTTGTAGGATG 124
QY 120 CTTCCAGTCAGGTGAGGTGTACAGGCCCAAGAGCAACATAATCTCTAGAACAAACA 179
Db 125 CTTCCAGTCAGGTGAGGTGTACAGGCCCAAGAGCAACATAATCTCTAGAACAAACA 181
QY 180 AGATCCCTTCAACTCTCTCTCATATCTTTGGTGTACAGGTGATCCCAATGGTGGACAAAT 239
Db 182 CCCTCCCT 238
QY 240 AGGAGTAACTACTATACCTTTAAATCTTCAACCACTCTCTCTCTCTCTCTCTCTCTCTCT 299
Db 239 TGGTCCAGGAAATTTGAGAGCAACCCCTCCACCCCTTTGGTCCAGGAATTTGGTCCAGCACC 298
QY 300 CGGTTTTATCACTGGACCACTATTTGGTGTACAGGTACCACTGAATATCAATATCAGTG 359
Db 299 CCCTCCACCT 341
QY 360 GCAGCTAACTGCTCCAGACCCCTACACCTCTAAGCAATCTTCAACACTACTCGGACTACCCAAAA 479
Db 342 -CCAAGCAATCCGCT 400
QY 420 AGAACAGCAAAATACAAAACAGATGCCAAATCTTCAACACTACTCGGACTACCCAAAA 479
Db 401 AGTACAGCAACTACTACTGCGCAGTGCAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 460
QY 480 TTCCACTGATATTTTGAAGGTGGTGGCAA 509
Db 461 TTCCACTGATATTTTGAAGGTGGTGGAA 490

RESULT 15
BF536022 653 bp mRNA linear EST 11-DEC-2000
LOCUS 602054151F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4193269 5',
DEFINITION mRNA sequence.
ACCESSION BF536022
VERSION BF536022.1 GI:11623390
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 653)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9525 row: e column: 14
High quality sequence stop: 653.
Location/Qualifiers
1. .653
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"

FEATURES
source

/db_xref="taxon:10090"
/clone="IMAGE:4193269"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Not1; Site 2: SalI; cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 28.9%; Score 147.2; DB 10; Length 653;
Best Local Similarity 61.5%; Pred. No. 2.4e-27;
Matches 313; Conservative 0; Mismatches 171; Indels 25; Gaps 4;
QY 2 AACTCACTGACCCAGAGAGCTTGTACCCAGCAGCATTTCCCGGTCTAGAGTTTCTCCAGG 61
Db |||||
1 AAATAACTGACCAGACCGCTTCTGACCAGCACATTCACCAGCTAAGAGTGTCTCCAAGA 60
QY 62 GGCT-ACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGGATCCTTCTAGCATGC 120
Db |||||
61 GGCTGAAAAAAGATGAAGCACTCAATTTGGTCTTAGGCCTCTGCATTTCTTAGGATGC 120
QY 121 TTCAGTCAAGTGTAGAGGCCCAAGAGACACATAATCTCTAGAAACAACAA 180
Db |||||
121 TTCTGTCTGTGAGTGTACAGAGGCCCTTAGAG--ACATGATCCTAGAGGCCATTC 177
QY 181 GATCCTTCAACTCTTCTCATTTATCTTGTCTTTCAGCCGTATCCCAATGGTGGACAAATA 240
Db |||||
178 CCTCCTCTCTCCCGCTCCCGCATGGTCCAGG-----AANTGGTAGACCACAC 222
QY 241 GGAGTAACAATCACTATACCTTAAATCTTCAACACCTCGTGTCTTGTGTAATCTTCCC 300
Db |||||
223 CCTCACCCTNTGGTCCAGGAATGGTAGACACCCCTCCACCCCTGTGTCCAGGAAT 282
QY 301 GGTTTTATCACTGGACCACTTGGTTGTACAGGTACCACTGAATATCAATATCAGTGG 360
Db |||||
283 GGTGACCAACCCCTCCAGCTCCTGTGCCACCAAGTTCCTCCACATCCTAGACCTCCAAG 342
QY 361 CAGCTAACTGCTCCAGACCCCTACACCTTAAGCAATCTCTCTACTCACTTCAATTCACA 420
Db |||||
343 AATCCGCTCTCCTCA-----CCTACTCCAAGCATTCCTCTACCGGACCTCTTACCACA 396
QY 421 GACAAAGCAATACAAAAACAGATGCCAAATCTCCAACACTACTGCGACTACCCAAAAT 480
Db |||||
397 GTACAGCAACTACATGCCAGCTGCCAGTATCTCCATACTACTCTCTACTGCTAGAGAT 456
QY 481 TCCACTGATATTTTCAAGGTGTGGCAA 509
Db |||||
457 TCCACTGATATTTTGGAGGTGTGGGAA 485

Search completed: July 30, 2004, 19:06:01
Job time : 2915 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 17:10:46 ; Search time 303 Seconds
(without alignments)
934.076 Million cell updates/sec

Title: US-09-386-850-7_COPY_1_510

Perfect score: 510

Sequence: 1 AAACGACTGACACAGAGC.....TTTTGAAGGTGTGCAAA 510

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	658	3	US-08-476-120-7
2	36	7.1	644	4	US-09-636-215-752
3	36	7.1	644	4	US-09-685-166A-752
4	35.8	7.0	832	4	US-09-621-976-2813
5	35	6.9	864	4	US-09-328-352-2584
6	35	6.9	7218	1	US-08-232-463-14
7	34.2	6.7	1150	4	US-09-461-325-69
8	34.2	6.7	1150	4	US-10-012-542-69
9	33.8	6.6	116592	4	US-09-818-512-3
10	33.6	6.6	5208	3	US-04-781-891-70
11	33.6	6.6	5208	4	US-09-791-211-11
12	33.6	6.6	5208	4	US-09-618-166-70
13	33.6	6.6	1664976	4	US-08-916-421B-1
14	33.4	6.5	528	4	US-09-105-542A-12
15	33	6.5	33	4	US-08-801-405B-8
16	32.8	6.4	399	4	US-09-621-976-8976
17	32.6	6.4	307	4	US-09-313-294A-6950
18	32.6	6.4	169938	4	US-09-676-610B-24
19	32.6	6.4	197496	4	US-09-877-177A-10
20	32.4	6.4	399	4	US-09-621-976-15653
21	32.4	6.4	425	4	US-09-621-976-15651
22	32.2	6.3	1193	4	US-09-461-325-119
23	32.2	6.3	1193	4	US-10-012-542-119
24	32.2	6.3	392000	4	US-10-027-983-11
25	32.2	6.3	1230025	4	US-09-198-452A-1
26	32	6.3	1333	3	US-08-889-425-3
27	32	6.3	3100	1	US-08-296-362-1

C	28	32	6.3	10144	4	US-10-204-708-94	Sequence 94, Appl
	29	31.8	6.2	809	3	US-09-276-531-27	Sequence 27, Appl
	30	31.8	6.2	2207	3	US-08-956-322-3	Sequence 3, Appl
	31	31.8	6.2	2329	3	US-08-956-322-1	Sequence 1, Appl
	32	31.8	6.2	3350	3	US-09-110-116-2	Sequence 2, Appl
	33	31.6	6.2	19233	4	US-10-204-708-45	Sequence 45, Appl
	34	31.4	6.2	1024	4	US-09-328-475C-20	Sequence 20, Appl
	35	31.4	6.2	1830121	4	US-09-557-884-1	Sequence 1, Appl
	36	31.4	6.2	1830121	4	US-09-643-990A-1	Sequence 1, Appl
	37	31.2	6.1	3211	2	US-08-574-959A-8	Sequence 8, Appl
	38	31.2	6.1	3211	3	US-09-357-014-8	Sequence 8, Appl
	39	31.2	6.1	3901	2	US-08-574-959A-6	Sequence 6, Appl
	40	31.2	6.1	3901	3	US-09-357-014-6	Sequence 6, Appl
	41	31	6.1	678	4	US-09-328-352-713	Sequence 713, App
	42	31	6.1	891	4	US-09-134-001C-139	Sequence 139, App
	43	31	6.1	5163	3	US-08-700-651-1	Sequence 1, Appl
	44	31	6.1	5163	3	US-08-928-361B-4	Sequence 4, Appl
	45	31	6.1	5163	4	US-09-588-995A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-476-120-7
Sequence 7, Application US/08476120
Patent No. 6025143

GENERAL INFORMATION:

APPLICANT: Rosinski-Chupin, Isabelle R.

APPLICANT: Tronik, Diana

APPLICANT: Rougeon, Francois

APPLICANT: Seidah, Nabil

TITLE OF INVENTION: Peptides and Polypeptides Derived

TITLE OF INVENTION: from the Submaxillary Gland of the Rat, Corresponding

TITLE OF INVENTION: Antibodies, Corresponding Hybridomas, and Uses of

TITLE OF INVENTION: These Products for Diagnosis, Detection, or

TITLE OF INVENTION: Therapeutic Purposes

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,120

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/153,277

FILING DATE: 17-NOV-1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/499,276

FILING DATE: 19-JUL-1990

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR88/00523

FILING DATE: 11-OCT-1989

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 88/13353

FILING DATE: 11-OCT-1988

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 004900-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Rat
TISSUE TYPE: Submaxillary Gland
CELL TYPE: Glandular
POSITION IN GENOME:
CHROMOSOME/SEGMENT: SMR1
FEATURE:
NAME/KEY: CDS
LOCATION: 73...510
US-08-476-120-7

Query Match 100.0%; Score 510; DB 3; Length 658;
Best Local Similarity 100.0%; Pred. No. 3.4e-146;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATGACTGACGAGAGCTTCTGACGAGCATTTCCCGCTCAGAGATTCTCCAG 60
DB 1 AAATGACTGACGAGAGCTTCTGACGAGCATTTCCCGCTCAGAGATTCTCCAG 60

QY 61 GGGCTACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGATCTCTTAGCATGC 120
DB 61 GGGCTACCAAGATGAAGTCACTGATTTGATCTTTGGCTGTGATCTCTTAGCATGC 120

QY 121 TTCCAGTCAGGTGAGGGTGTGAGAGCCCAAGAGACAAATATCTTAGAGACAA 180
DB 121 TTCCAGTCAGGTGAGGGTGTGAGAGCCCAAGAGACAAATATCTTAGAGACAA 180

QY 181 GATCTTCAACTTCTCCCTCATTTATCTTGCTCTTCAGCCTGATCCCAATGGTGACAA 240
DB 181 GATCTTCAACTTCTCCCTCATTTATCTTGCTCTTCAGCCTGATCCCAATGGTGACAA 240

QY 241 GGAGTAACAATCACTATACCTTAATCTTCAACCACTCGTGTCTTCTGTTAATCTTCCC 300
DB 241 GGAGTAACAATCACTATACCTTAATCTTCAACCACTCGTGTCTTCTGTTAATCTTCCC 300

QY 301 GGTGTTATCACTGGACCACTGTTGTTGACAGGTACCACTGAATATCAATATCACTGG 360
DB 301 GGTGTTATCACTGGACCACTGTTGTTGACAGGTACCACTGAATATCAATATCACTGG 360

QY 361 CAGCTAACTGCTCCAGACCTCACTCACTTAAGCAATCTCTCACTCACTTCACTTCCACA 420
DB 361 CAGCTAACTGCTCCAGACCTCACTCACTTAAGCAATCTCTCACTCACTTCACTTCCACA 420

QY 421 GAACAAGCAATACAAAAACAGATGCGCAAAATCTCCAACTACTCGAGTACCCAAAAT 480
DB 421 GAACAAGCAATACAAAAACAGATGCGCAAAATCTCCAACTACTCGAGTACCCAAAAT 480

QY 481 TCCACTGATATTTTGAAGTGGTGCAAA 510
DB 481 TCCACTGATATTTTGAAGTGGTGCAAA 510

RESULT 2
US-09-636-215-752/c
Sequence 752, Application US/09636215
Patent No. 6620922
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi

APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.42717C21
CURRENT APPLICATION NUMBER: US/09/636,215
NUMBER OF SEQ ID NOS: 852
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 752
LENGTH: 644
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(644)
OTHER INFORMATION: n=A,T,C or G
US-09-636-215-752

Query Match 7.1%; Score 36; DB 4; Length 644;
Best Local Similarity 58.3%; Pred. No. 0.23;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 384 ACCTTAAGCAATCTCTTACTCACTTCACTTCCACGACCAAGCAAAATACAAAACAGA 443
DB 126 ACCTCAAGGCAAACTCAACTCAAAAGGGAACCCGACCAAGGACACACCAAC 67

QY 444 TCCCAAAATCTCCAACTACTTGGCACTACCCAAAATTCCTGATAT 491
DB 66 TCGCAAGCACTCATACCAAGTTCTTACGACCCCAAAAATCCCAAT 19

RESULT 3
US-09-685-166A-752/c
Sequence 752, Application US/09685166A
Patent No. 6630305
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C21
CURRENT APPLICATION NUMBER: US/09/685,166A
NUMBER OF SEQ ID NOS: 898
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 752
LENGTH: 644

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(644)
; OTHER INFORMATION: n=A,T,C or G
US-09-685-166A-752

Query Match
Best Local Similarity 7.1%; Score 36; DB 4; Length 644;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 384 ACCTCTAAGCAATCCTCTACTCACTCACTTCCACAGAACAGCAATACAAACAGA 443
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126 ACCTCAAGGCAAACTCACTCAAAAGGAACCCGACGAGGACACCAACAAC 67
QY 444 TGCAGAAATCTCCAACTACTCGGACTACCCAAATTCACACTGATAT 491
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 TCGCAAGACCTCATACCAGGTCTCTACGACCCCAAAATCCACCCAAT 19

RESULT 4
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 7.0%; Score 35.8; DB 4; Length 832;
Matches 31; Conservative 115; Mismatches 107; Indels 0; Gaps 0;

QY 186 TTCAACTCTCTCTCATTTATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATAGGAGT 245
Db :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : ::::
49 WKYKWTWYWRFYAMWGYKKAMCRTKTKKKKKGYWMMYWGWRYSYMAWTRWTG 108
QY 246 AACAATCACTATACCTTTAAATCTTCAACACCTCGTGTCTTGTAAATCTTCCCGGTTT 305
Db :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : ::::
109 YAYRSMYTWYRYRCWKKKAYRKTTCYSSKGTWWRKWKAWTTWWKKTYWAATRYW 168
QY 306 TATCACTGACCAACCATTTGGTCTTCAAGGTACCACTGATATATCAATATCAGTGGCAGCT 365
Db :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : ::::
169 WMCWTKRWASWYWCWWGKARKWTKWRSYSASAKRCCYSCSWGMSWKYMMRW 228
QY 366 AACTGCTCCAGACCCCTACACCTTAAGCAATCTCTCACTCACTTCAATCCACAGACA 425
Db :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: : ::::
229 WRWRGWATGAGMKAWRASCMWRKYSKTSYKSWMMCWTRSKYCYTKARTGYCYR 288
QY 426 AGCAATACAAAA 438
Db ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
289 KGGMWKGRGRWYA 301

RESULT 5
US-09-328-352-2584
; Sequence 2584, Application US/09328352
; Patent No. 6562958
```

```
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2584
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2584

Query Match
Best Local Similarity 6.9%; Score 35; DB 4; Length 864;
Matches 98; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 104 GGATCCTTTCAGCATGCTTCCAGTCAGGTGAGGTGTTCAGAGGCCCAAGAGACAACATA 163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
397 GAATTTTATATGAAGCTTTTACTGAAATGAATGAGACATTATTGTATGCGCAAAATTA 456
QY 164 ATCCTAGAGACACACAGATCCTTCAACTTTCCTCATTTCTTGGTCTTCAGCCTGATC 223
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
457 ATTTTGAAGAGAGAGAGCTCTTCTCCATTCCTGATTCATTTGTTATTACCTCATTTT 516
QY 224 CCAATGGTGACAAATAGGAGTAACAATCACTATACCTTAAATCTTCAACCACTCGTG 283
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
517 ATTTAGAGAAATACAAAAGCCTCAATATTATATAAGACTTTTAAACCTGATCGAT 576
QY 284 TTCTTGTAAATCTTCCCGGTTT 306
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
577 GTGATTTTCATCTCTCAGAGGGTTT 599

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
```

Db	876	GAGGTGTAACGACAAAGAGAGATTTTAAAGTTCAGAAATGGTTACAGAAGTATTAAGACAGCT	935
Qy	75	GAAGTCACCTGATTTTCATCTTTGGCCGTGGATCCCTTCTAGCATGCTTCAGATCAGGTGA	134
Db	936	GTCTGGGTGTTTTTTCGATTTTGTGGTTCTCGGTTTCAATCTCTGCTCATTTCAACAAGATGG	995
Qy	135	GGGTGTTCAGAGGCCCAAGAGACCAACATAATCTAGAAGACCAACAAGATCCTTCAACTCT	194
Db	996	GAGTTTTATAGAACTAAAGACCATTGAAGCTACTTAAAAACAACAACAAAAAGGCTCA	1055
Qy	195	TCCTCAATTCTGGTCTTCAGCCTGATCCCAATG	229
Db	1056	TCAATTTCTCAGTCTGAATTCACAAAAATGCCAATG	1090

```

RESULT 8
US-10-012-542-69
; Sequence 69, Application US/10012542
; Patent NO. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-12-14
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-06-16
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-06-16
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-06-16
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-06-22
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-06-22
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-012-542-69

```

RESULT 9
US-09-818-512-3/c
; Sequence 3, Application US/09818512
; Patent No. 6537780

```

;
; TELEX: 899149
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9PT-F1s
;
US-08-232-463-14

Query Match          6.9%; Score 35; DB 1; Length 7218;
Best Local Similarity 6.0%; Pred. No. 1.4;
Matches 17; Conservative 148; Mismatches 118; Indels 0; Gaps 0;

Qy      183 TCCTTCAACTCTTCCTCATATATCTGTCTTCAGCCCTGATCCCAATGGTGGACAAATAGG 242
Db      1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241

Qy      243 AGTAACAATCAGTATACCCCTTAATCTTCAACACCTCGTCTTCTGTTCTTAATCTTCCCG 302
Db      1242 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1301

Qy      303 TTTTATCACTGGACCAACCATGGTTGTACAAGGTACCACCTCAATATCAATATCAGTGGCA 362
Db      1302 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1361

Qy      363 GCTAACTGTCTCAGACCTACACTCTTAAGCAATCTCTTACTCAACTTCATTCACAGA 422
Db      1362 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1421

Qy      423 ACAAGCAAAATACAAAACAGATGCGAAAATCTCCAACTACT 465
Db      1422 YYYYYYYYYYYYGTACCAAACTCTTCTATCTCTTAACTACT 1464

```

```

RESULT 7
US-09-461-325-69
; Sequence 69, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461.325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-325-69

```

Query Match	6.7%;	Score 34.2;	DB 4;	Length 1150;
Best Local Similarity	47.4%;	Pred. No. 1.1;		
Matches 102;	Conservative	0;	Mismatches 113;	Indels 0; Gaps 0

```
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
; US-09-818-512-3

Query Match      6.6%; Score 33.8; DB 4; Length 116592;
Best Local Similarity 62.4%; Pred. No. 12;
Matches 53; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 367 ACTGCTCCAGACCTACACCTCTAAGCAATCCCTACTCAACTTCATTCACAGAACAA 426
Db 1617 ACTGCACTCAGCTAGGCAACAGCAAGACTCTGTCTCAAAAAACAAACAAACAA 1558

QY 427 GCATATCAAAAAACAGATGCCAAAA 451
Db 1557 AACAAAAACAAAAACAAATGAGTAAA 1533

RESULT 10
US-08-781-891-70/c
; Sequence 70, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5208 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-781-891-70

Query Match      6.6%; Score 33.6; DB 3; Length 5208;
Best Local Similarity 59.4%; Pred. No. 3.3;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 329 TACAAGGTACCACTGAATATCAATATCAATGAGTGGCAGCTAACTGCTCCAGACCCCTACACCTC 388
Db 1623 TACATAGGACGTATCGTTTTCATATCATTTGGAGATAATGCTTAAGCATCTCCATTTC 1564

QY 389 TAAGCAATCTCTACTCAACTTCATTCACAGAAC 424
Db 1563 TAAATCTTCATCTCTCAATTACATAGGACGTATC 1528

RESULT 11
US-09-791-211-11/c
; Sequence 11, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 11
; LENGTH: 5208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(4530)
; US-09-791-211-11

Query Match      6.6%; Score 33.6; DB 4; Length 5208;
Best Local Similarity 59.4%; Pred. No. 3.3;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 329 TACAAGGTACCACTGAATATCAATATCAATGAGTGGCAGCTAACTGCTCCAGACCCCTACACCTC 388
Db 1623 TACATAGGACGTATCGTTTTCATATCATTTGGAGATAATGCTTAAGCATCTCCATTTC 1564

QY 389 TAAGCAATCTCTACTCAACTTCATTCACAGAAC 424
Db 1563 TAAATCTTCATCTCTCAATTACATAGGACGTATC 1528

RESULT 12
US-09-618-166-70/c
; Sequence 70, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/618,166
  FILING DATE: 17-Jul-2000
  CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Mcmasters, David D.
  REGISTRATION NUMBER: 33,963
  REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (206) 622-4900
  TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 70:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 5208 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-618-166-70
Query Match          6.6%; Score 33.6; DB 4; Length 5208;
Best Local Similarity 59.4%; Pred. No. 3.3;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 329 TACAAGGTACCACTGAATATCAATATCACTGCGACGTAAGTCTCCAGACCCCTACACCTC 388
Db 1623 TACATAGGACGTATCGTTTTCATTATCATTTGGAGATAAATGCTTAAGCATCTCCATTTC 1564
QY 389 TAAGCAATCCCTCACTCACTCACTTCATCCACAGAAC 424
Db 1563 TAAATCTTCACACTCTCAATTACATAGGACGTATC 1528

RESULT 13
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
;
; LOCATION: (98120)..(98120)
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 6.6%; Score 33.6; DB 4; Length 1664976;
Best Local Similarity 57.7%; Pred. No. 43;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 363 GCTAACTGCTCCAGACCCCTACACCTCTAAGCAATCTCTTACTCACTCACTTCACTTCCACAGA 422
Db 1364462 GTTATCAACTCCAGTTCCTACAGCAGCAATAATCCCTGCTATTGGAAGTAAATCTACTT 1364521
QY 423 ACAAGCAATACAAAACAGATGCCAAATCTTCAACACTACTG 466
Db 1364522 CCATCTATTAAAGAGCAAGCCCAATATAATAAATCTCTG 1364565

RESULT 14
US-09-105-542A-12
; Sequence 12, Application US/09105542A
; Patent No. 6323329
; GENERAL INFORMATION:
; APPLICANT: Bullerdiel, Jörn
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES OF GENES OF THE
; FILE REFERENCE: BOEHA 001C1CP
; CURRENT APPLICATION NUMBER: US/09/105,542A
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: PCT/DE96/02494
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(528)
; OTHER INFORMATION: n = A,T,C or G
US-09-105-542A-12

Query Match 6.5%; Score 33.4; DB 4; Length 528;
Best Local Similarity 60.4%; Pred. No. 1.3;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 420 AGAACAAGCAATACAAAACAGATGCCAAATCTTCAACACTACTGCGACTACCCAAA 479
Db 241 AGGAATGGAATACAAAATACATCTCAAAATCTGTAATAATCTGAAGGACCTCTA 300
QY 480 TTCCACTGATATTTTGAAGTGGGCAAA 510
Db 301 TGGCCAAAATAATCTTGAAGAAGATGAAAAA 331

RESULT 15
US-08-801-405B-8
; Sequence 8, Application US/08801405B
; Patent No. 6589750
; GENERAL INFORMATION:
; APPLICANT: ROUGEOT, Catherine
; TITLE OF INVENTION: THERAPEUTIC USE OF THE SMRI PROTEIN, THE
; SMRI MATURATION PRODUCTS, SPECIFICALLY THE QHNP
; PENTAPEPTIDE AS WELL AS ITS BIOLOGICALLY ACTIVE
; DERIVATIVES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.405B

Mon Aug 2 07:39:40 2004

```
;
; FILING DATE: 20-Feb-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Dadio, Susan M.
; REGISTRATION NUMBER: 40,373
; REFERENCE/DOCKET NUMBER: 012880-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..33
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-801-405B-8
Query Match 6.5%; Score 33; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 139 GTCAGAGCCCCAAGACACATATCTCTAGA 171
Db 1 GTCAGAGCCCCAAGACACATATCTCTAGA 33
```

Search completed: July 30, 2004, 19:11:21
Job time : 313 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 13:32:45 ; Search time 838 Seconds
(without alignments)
2585.419 Million cell updates/sec

Title: US-09-386-850-7_COPY_1_510
Perfect score: 510
Sequence: 1 AAACGACTGACGAGAGAGC.....TTTTGAAGGTGGTGGCAA 510

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	658	2	AAQ04272
2	142	27.8	639	9	ADD29626
3	65	12.7	65	6	ABN30493
4	56.8	11.1	366	2	AAV89531
5	56.8	11.1	521	3	AAQ04378
6	53.4	10.5	153	6	ABK55154
7	47.2	9.3	1956	3	AAQ04378
8	43.6	8.5	1362	3	AAQ04908
9	43	8.4	610	2	AAV89871
10	41.2	8.1	5883	9	ADB54310
11	41	8.0	5085	7	ACF62813
12	41	8.0	8085	9	ADB54279
13	39.4	7.7	65	6	ABN56280
14	38.2	7.5	56153	3	AAQ46793
15	38.2	7.5	1391	4	AAQ55632
16	38.2	7.5	6022	4	AAQ46661
17	38.2	7.5	40304	8	ADA03014
18	38.2	7.5	40304	9	ADB72752
19	38.2	7.5	40304	9	ADC85494
20	38.2	7.5	86574	6	ABK83560
21	38	7.5	2000	7	ADA71938
22	37.8	7.4	495	6	ABN71324
23	37.8	7.4	1870	7	ACC51126

24	37.8	7.4	110000	6	ABN71527_11
25	37.8	7.4	110000	6	ABN71527_12
26	37.4	7.3	19653	6	ABL33335
27	37	7.3	105325	6	ABK94407
28	36.4	7.1	5127	4	AAQ45425
29	36.4	7.1	6222	6	ABL32693
30	36.2	7.1	5815	6	ABK40023
31	36.2	7.1	8776	6	ABK40068
32	36.2	7.1	110000	7	ABX16390_2
33	36	7.1	644	4	AAQ46085
34	36	7.1	644	5	ACA59893
35	36	7.1	644	6	ABL5456
36	36	7.1	644	7	ACC95620
37	36	7.1	644	9	ADB14202
38	35.8	7.0	604	7	ABX98613
39	35.8	7.0	7029	6	ABX2620
40	35.8	7.0	7029	6	AAQ28384
41	35.6	7.0	549	8	ACC73187
42	35.6	7.0	4255	6	ABL34033
43	35.2	6.9	414	5	ABV01994
44	35.2	6.9	886	4	AAI94116
45	35	6.9	864	8	ADA31297

ALIGNMENTS

RESULT 1
AAQ04272
ID AAQ04272 standard; DNA; 658 BP.
AC AAQ04272;
XX
DT 21-SEP-1989 (first entry)
XX
DE SMR1 encoding gene.
XX
KW Submaxillary gland; behavioural control; ss.
XX
OS Rattus.
FH Key
FT Location/Qualifiers
FT CDS
FT 73..510
FT /*tag= a
XX
PN WO9003981-A.
XX
PD 19-APR-1990.
XX
PF 11-OCT-1988; 88FR-00013353.
XX
PR 11-OCT-1988; 88FR-00013353.
XX
PA (INSP) INST PASTEUR.
XX
PI Chupin I, Tronik D, Rougeon F, Seidah N;
XX
DR WPI; 1990-147823/19.
XX
PT New polypeptide useful therapeutically and in diagnosis - isolated from
PT rat submaxillary gland and derived tetra: or penta:peptide(s), antibodies
PT and hydridomas.
XX
PS Disclosure; Page ?; 26pp; French.
XX
CC This sequence encodes the SMR1 polypeptide which is secreted from rodent
CC esp. rat submaxillary glands and is associated with control of behaviour
CC in such animals. Abs derived from SMR1 are useful for detecting its
CC presence in biological tissues and fluids. See also AAR04389 and AAR04420
CC -25
XX
SQ Sequence 658 BP; 216 A; 162 C; 111 G; 169 T; 0 U; 0 Other;

Continuation (12 o
Continuation (13 o
Abl33335 Human imm
Abk94407 DNA encod
Aas45425 Chemical
Abl32693 Human imm
Abk40023 Human che
Abk40068 Human che
Continuation (3 of
Aas64085 Human pro
Aas59893 Prostate
Abl95456 Human pro
Acc95620 Prostate
Abl14202 Human pro
Abx98613 Rice leaf
Abl32620 Human imm
Aad28384 Human che
Acc73187 Cat flea
Abl34033 Human imm
Abv01994 Human pro
Aai94116 Human neu
Ada31297 DNA encod

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
PR
XX 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
PA
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX

XX Example 1; SEQ ID NO 3241; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 65 BP; 12 A; 18 C; 12 G; 23 T; 0 U; 0 Other;
Query Match 12.7%; Score 65; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.6e-10;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 277 CCTCGTGTCTGTAAATCTTCCCGTTTATCACTGGACCACTTGTGTACAAGT 336
|||||
Db 1 CCTCGTGTCTGTAAATCTTCCCGTTTATCACTGGACCACTTGTGTACAAGT 60
|||||

QY 337 ACCAC 341
|||||
Db 61 ACCAC 65

RESULT 4
AAV89531
ID AAV89531 standard; cDNA; 366 BP.
XX
AC AAV89531;
XX

DT 15-FEB-1999 (first entry)
XX DE EST clone CP314.
XX
KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
XX WO9845436-A2.
PN
XX 15-OCT-1998.
PD
XX 10-APR-1998; 98WO-US006955.
XX
XX 10-APR-1997; 97US-00838821.
PR
XX (GEMY) GENETICS INST INC.
PA
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX WPI; 1999-070077/06.
DR

XX New polynucleotides encoding human secreted proteins - derived from e.g.
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.
XX

PS Claim 1; Page 247; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST). The
CC polynucleotide, which is a secreted EST, and the encoded protein are
CC predicted to have useful biological activities which would make them
CC suitable for treating, preventing or ameliorating medical conditions in
CC humans and animals, although no supporting data is given. Suggested
CC activities include nutritional activity, immune stimulating or
CC suppressing activity, haematopoiesis regulating activity, tissue growth
CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-
CC inflammatory activity, cadherin/tumour invasion suppressor activity,
CC tumour inhibition activity. The polynucleotide may also be useful for
CC gene therapy
XX

XX Sequence 366 BP; 100 A; 100 C; 85 G; 81 T; 0 U; 0 Other;

Query Match 11.1%; Score 56.8; DB 2; Length 366;
Best Local Similarity 60.3%; Pred. No. 5e-07;
Matches 94; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 69 AAGATGAAGTCACCTGTATTGATCTTTGGCTGTGGATCCTTCTAGCATGCTTCAGTC 128
|||||

Db 80 AAGGATGAATCCTGACTTGTGATCTTTGGCTGTGGCTCTTGCAGCGTGTTCACACC 139
|||||

QY 129 AGGTGAGGCTGTCAAGGCCCAAGAGACAACTAATCTTAGAACAACAAGATCCTTC 188
|||||

Db 140 TGGTGAGAGTCAAGAGGCCCAAGAGACAACTAATCTTAGAACAACAAGATCCTTC 199
|||||

QY 189 AACTCTTCTCATATCTTGTCTTCACCTGTATCC 224
|||||

Db 200 ACCCTTTGGCCAGGATTTGTTCCACCACCTCTCC 235
|||||

RESULT 5
AAC04378
ID AAC04378 standard; cDNA; 521 BP.
XX
AC AAC04378;
XX

DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 8453.

XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.	XX	WO200212280-A2.
XX	Homo sapiens.	XX	14-FEB-2002.
XX	EPI033401-A2.	XX	30-JUL-2001; 2001WO-US023826.
XX	06-SEP-2000.	XX	03-AUG-2000; 2000US-0223265P.
XX	21-FEB-2000; 2000EP-00200610.	XX	02-OCT-2000; 2000US-0237406P.
XX	26-FEB-1999; 99US-0122487P.	XX	20-MAR-2001; 2001US-0277495P.
XX	(GBST) GENSET.	XX	03-JUL-2001; 2001US-0302702P.
XX	Dumas Milne Edwards J, Duclert A, Giordano J;	XX	(CORI-) CORIXA CORP.
XX	WPI; 2000-500381/45.	XX	Pyle RA, Xu J, Secretist H;
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for	XX	WPI; 2002-257462/30.
XX	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for	XX	Novel polynucleotide encoding colon tumor polypeptides, useful as
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures.	XX	vaccines for treating colon cancers.
XX	Claim 1; SEQ ID NO 8453; 71pp + Sequence Listing; English.	XX	Claim 1; Page 327; 425pp; English.
XX	The present sequence is one of a large number of 5' ESTs derived from	XX	The invention relates to isolated polynucleotides (I) encoding colon
XX	mRNAs encoding secreted proteins. No ORF has yet been conclusively	XX	tumour polypeptides (II). (I) is useful for stimulating an immune
XX	identified within the present sequence. The 5' ESTs were prepared from	XX	response in a patient and treating colon cancer in a patient.
XX	total human RNAs or polyA+ RNAs derived from 30 different tissues. EST	XX	Oligonucleotides derived from (I) are useful for determining the presence
XX	sequences usually correspond mainly to the 3' untranslated region (UTR)	XX	of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX	of the mRNA because they are often obtained from oligo-dT primed cDNA	XX	compositions, e.g. vaccines, and other compositions for the diagnosis and
XX	libraries. Such ESTs are not well suited for isolating cDNA sequences	XX	treatment of colon cancer. A composition comprising a first component
XX	derived from the 5' ends of mRNAs and even in those cases where longer	XX	selected from physiologically acceptable carriers and immunostimulants,
XX	cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'	XX	and an antigen-presenting cell expressing (II) is useful for inhibiting
XX	ESTs are derived from mRNAs with intact 5' ends and can therefore be used	XX	development of cancer in a patient. (I) is useful in the design and
XX	to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in	XX	preparation of ribozyme molecules for inhibiting expression of tumour
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures.	XX	polypeptides and (II). ABK54531-ABK55464 represent human colon cancer cDNA
XX	They are used to obtain upstream regulatory sequences and to design	XX	sequences of the invention
XX	expression and secretion vectors	XX	Sequence 153 BP; 36 A; 43 C; 39 G; 35 T; 0 U; 0 Other;
XX	Sequence 521 BP; 122 A; 187 C; 87 G; 123 T; 0 U; 2 Other;	XX	Query Match 10.5%; Score 53.4; DB 6; Length 153;
XX	Query Match 11.1%; Score 56.8; DB 3; Length 521;	XX	Best Local Similarity 63.8%; Pred. No. 3.7e-06;
XX	Matches 94; Conservative 0; Mismatches 62; Indels 0; Gaps 0;	XX	Matches 81; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY	69 AAGATGAAGTCACTGTAATTTGATCTTTGGCCTGTGGATCCTTCTAGCATGCTTCCAGTC 128	QY	69 AAGATGAAGTCACTGTAATTTGATCTTTGGCCTGTGGATCCTTCTAGCATGCTTCCAGTC 128
Db	100 AAGATGAATCACTGACTTGGATCTTGGCCTTTGGCTCTTGCAGCGTGTTCACACC 159	Db	26 AAGATGAATCACTGACTTGGATCTTGGCCTTTGGCTCTTGCAGCGTGTTCACACC 85
QY	129 AGGTGAGGTTGCAGAGGCCCAAGACACATATCTAGAGACACACAGATCCTTC 188	QY	129 AGGTGAGGTTGCAGAGGCCCAAGACACATATCTAGAGACACACAGATCCTTC 188
Db	160 TGGTGAGAGTCAAAAGAGGCCCAAGACACATATCACCCTGGACCGTGGCTCTCTCTCA 219	Db	86 TGGTGAGAGTCAAAAGAGGCCCAAGACACATATCACCCTGGACCGTGGCTCTCTCTCA 145
QY	189 AACTCTTCCTATATCTTGGTCTTCAGCCTGATCC 224	QY	189 AACTCTT 195
Db	220 ACCTTTTGGCCAGGATTTGTTCCACCCTCTCC 255	Db	146 ACCTTTT 152
RESULT 6		RESULT 7	
ABK55154		AAC38547/c	
ID	ABK55154 standard; cDNA; 153 BP.	ID	AAC38547 standard; DNA; 1956 BP.
XX	ABK55154;	XX	AAC38547;
XX	18-JUN-2002 (first entry)	XX	17-OCT-2000 (first entry)
XX	Human colon cancer-associated cDNA, SEQ ID No 624.	XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 21366.
XX	Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.	XX	Hybridisation assay; genetic mapping; gene expression control;
XX	Homo sapiens.	XX	protein identification; signal transduction pathway; metabolic pathway;
OS		XX	promoter; termination sequence; ss.
		OS	Arabidopsis thaliana.
		XX	EP1033405-A2.

XX PD 06-SEP-2000.
XX PD 25-FEB-2000; 2000EP-00301439.
XX PD 25-FEB-1999; 99US-0121825P.
XX PD 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014287D.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

PR	22-SEP-1999;	99US-01551139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157533P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	

Query Match

Best Local Similarity 9.3%; Score 47.2; DB 3; Length 1956;

Pred. No. 0.001;

Mismatches 127; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY	157	CRACATAATCTAGAGACAAGATCCTTCAACTCTTCTCATTAATCTTGGTCTTCAG	216
Db	843	CAGCAAGCATCAAACTCAATCTTCTTCAATCAATCAATCTTCTTCTTCAT	784
QY	217	CCTGATCCCAATGGTGGCAAAATAGAGTAACAATCACTATACCCCTTAATCTTCAACCA	276
Db	783	CATCGCTCAAAATTIGCTCTTCTCTGTGCCACTCTTTTGTGTTTGTGCTATCC	724
QY	277	CCTCGTGTCTTGTTAATCTTCCGGTCTTATCACTGGACCACCATTTGTGTACAGGT	336
Db	723	CATACTCTTCTGCTTCCACTCTCTTAAATAATCTTCAATCTTCTTCTTGAAGAACT	664
QY	337	ACCACATGAATATCAATATCAGTGGCAGCTAACTGCTCCAGACCTCAACCTCTTAAGCAAT	396
Db	663	TATCCTCTATCTCTTCAATGTCTCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT	604
QY	397	CCTCCTACTCAACTTCATTC	416
Db	603	CTTCT	584

RESULT 8

AAC49080/c

ID AAC49080 standard; DNA; 1362 BP.

XX

AC AAC49080;

XX

XX 18-OCT-2000 (first entry)

DT

XX

DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 59850.
XX	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
KW	promoter; termination sequence; ss.
XX	Arabidopsis thaliana.
OS	Arabidopsis thaliana.
XX	EP1033405-A2.
PN	06-SEP-2000.
XX	25-FEB-2000; 2000EP-00301439.
PF	25-FEB-1999;
XX	99US-0121825P.
PR	05-MAR-1999;
PR	99US-0123180P.
PR	09-MAR-1999;
PR	99US-012348P.
PR	23-MAR-1999;
PR	99US-0125788P.
PR	25-MAR-1999;
PR	99US-0126264P.
PR	29-MAR-1999;
PR	99US-0126785P.
PR	01-APR-1999;
PR	99US-0127462P.
PR	06-APR-1999;
PR	99US-0128234P.
PR	08-APR-1999;
PR	99US-0128714P.
PR	16-APR-1999;
PR	99US-0129845P.
PR	19-APR-1999;
PR	99US-0130077P.
PR	21-APR-1999;
PR	99US-0130449P.
PR	23-APR-1999;
PR	99US-0130510P.
PR	28-APR-1999;
PR	99US-0130891P.
PR	30-APR-1999;
PR	99US-0131449P.
PR	30-APR-1999;
PR	99US-0132048P.
PR	30-APR-1999;
PR	99US-0132407P.
PR	04-MAY-1999;
PR	99US-0132484P.
PR	05-MAY-1999;
PR	99US-0132485P.
PR	06-MAY-1999;
PR	99US-0132487P.
PR	07-MAY-1999;
PR	99US-0132863P.
PR	11-MAY-1999;
PR	99US-0134256P.
PR	14-MAY-1999;
PR	99US-0134218P.
PR	14-MAY-1999;
PR	99US-0134219P.
PR	14-MAY-1999;
PR	99US-0134221P.
PR	14-MAY-1999;
PR	99US-0134370P.
PR	18-MAY-1999;
PR	99US-0134768P.
PR	19-MAY-1999;
PR	99US-0134941P.
PR	20-MAY-1999;
PR	99US-0135124P.
PR	21-MAY-1999;
PR	99US-0135353P.
PR	24-MAY-1999;
PR	99US-0135629P.
PR	25-MAY-1999;
PR	99US-0136021P.
PR	27-MAY-1999;
PR	99US-0136392P.
PR	28-MAY-1999;
PR	99US-0136782P.
PR	01-JUN-1999;
PR	99US-0137222P.
PR	03-JUN-1999;
PR	99US-0137528P.
PR	04-JUN-1999;
PR	99US-0137502P.
PR	07-JUN-1999;
PR	99US-0137724P.
PR	08-JUN-1999;
PR	99US-0138094P.
PR	10-JUN-1999;
PR	99US-0138540P.
PR	10-JUN-1999;
PR	99US-0138847P.
PR	14-JUN-1999;
PR	99US-0139119P.
PR	16-JUN-1999;
PR	99US-0139452P.
PR	16-JUN-1999;
PR	99US-0139453P.
PR	17-JUN-1999;
PR	99US-0139492P.
PR	18-JUN-1999;
PR	99US-0139454P.
PR	18-JUN-1999;
PR	99US-0139455P.
PR	18-JUN-1999;
PR	99US-0139456P.
PR	18-JUN-1999;
PR	99US-0139457P.
PR	18-JUN-1999;
PR	99US-0139458P.
PR	18-JUN-1999;
PR	99US-0139459P.
PR	18-JUN-1999;
PR	99US-0139460P.
PR	18-JUN-1999;
PR	99US-0139461P.
PR	18-JUN-1999;
PR	99US-0139462P.
PR	18-JUN-1999;
PR	99US-0139463P.
PR	18-JUN-1999;
PR	99US-0139750P.
PR	18-JUN-1999;
PR	99US-0139763P.
PR	21-JUN-1999;
PR	99US-0139817P.

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145148P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 8.5%; Score 43.6; DB 3; Length 1362;
Best Local Similarity 48.1%; Pred. No. 0.011;
Matches 124; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 157 CAACATATCTTAGAGACAAAGATCTTCAACTCTTCTCATATATCTTGGTCTTCAAG 216
Db 481 CAGCAAAAGCATCAAACTCAACATCTTCTTCTCATCATCATCATCTTCTTCTTCTCAT
QY 217 CCTGATCCCAATGGTGGACAAATAGAGTAACAATCACTATACCTTAAATCTTCAACA 276
Db 421 CATCGCTCAATTTTGTCTTCTTCTGTGCCACTCTTTTGTGTTTGTGGTCTATCC 362
QY 277 CCTCGTGTCTTGTTPAATCTTCCCGTTTTATCACTGGACCACCACTGGTTGTACAAGT 336
Db 361 CATACTCTTCTGCTTCACTTCTCTTAAATAATCTTCTTCTTCTTCTTCTTCTTCTTCTT
QY 337 ACCACTGAATATCAATATCAGTGGAGCTAAGTGTCTCAGACCCCTACACCTCTTAAGCAAT 396
Db 301 TATCCTCTATTCTCTTCATTTGTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
QY 397 CCTCCTACTCAACTTCAT 414
Db 241 CTTCTTCTCTCTCTCTCTCT 224
```

RESULT 9
 ID AAV89871 standard; cDNA; 610 BP.
 XX
 AC AAV89871;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE EST clone CV915.
 XX
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO9845436-A2.
 XX
 PD 15-OCT-1998.
 XX
 XX 10-APR-1998; 98WO-US006955.
 PF
 XX 10-APR-1997; 97US-00838821.
 PR
 XX (GEMV) GENETICS INST INC.
 PA
 XX Jacobs K, McCooy JM, Lavallie ER, Racie LA, Merberg D, Treacy M,
 PI Spaulding V, Agostino MJ;
 PI
 XX WPI; 1999-070077/06.
 DR
 XX New polynucleotides encoding human secreted proteins - derived from e.g.
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
 PT pituitary, retina and colon cDNA libraries.
 PT
 XX Claim 1; Page 357; 618pp; English.
 PS
 XX The present sequence represents a human expressed sequence tag (EST). The
 CC polynucleotide, which is a secreted EST, and the encoded protein are
 CC predicted to have useful biological activities which would make them
 CC suitable for treating, preventing or ameliorating medical conditions in
 CC humans and animals, although no supporting data is given. Suggested
 CC activities include nutritional activity, immune stimulating or
 CC suppressing activity, haematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-
 CC inflammatory activity, cadherin/tumour invasion suppressor activity,
 CC tumour inhibition activity. The polynucleotide may also be useful for
 CC gene therapy
 XX
 SQ Sequence 610 BP; 169 A; 186 C; 84 G; 171 T; 0 U; 0 Other;
 Query Match 8.4%; Score 43; DB 2; Length 610;
 Best Local Similarity 56.8%; Pred. No. 0.011;
 Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 51 TTCTCAGAGGGCTACCAAGATGAAGTCACTGTTATTTGATCTTTGGCTGTGGATCCT 110
 Db 88 TTCTCAGAGGGCTACCAAGATGAAGTCACTGTTATTTGATCTTTGGCTGTGGATCCT 147
 QY 111 TCTAGCATGCTCCAGTCAGTGGGTGTCAGAGGCCAAGAGCAACAATATCTCTAG 170
 Db 148 TATTTCAATGTTTACACCCAGTGAAGTCAAGATTTCTCCAGAGCCATATCTACCTGG 207
 QY 171 AAGACAAACAGATCCTTTCA 189
 Db 208 CCAGTGCACCACTTCCA 226

RESULT 10
 ADB54310/c

ID ADB54310 standard; DNA; 5683 BP.
 XX
 AC ADB54310;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Pretreated genomic DNA region 234.
 XX
 KW colon cell proliferative disorder; non methylated CpG dinucleotide;
 KW cytosine; cancer; adenoma; carcinoma; cytosine methylation state; ds.
 XX
 OS Unidentified.
 XX
 PN WO2003072821-A2.
 XX
 PD 04-SEP-2003.
 XX
 XX 27-FEB-2003; 2003WO-EP002035.
 PF
 XX 27-FEB-2002; 2002EP-00004551.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;
 PI Rujan T, Schmitt A;
 PI
 XX WPI; 2003-731620/69.
 DR
 XX Detecting and differentiating between colon cell proliferative disorders
 PT associated with a gene or its regulatory regions comprises contacting a
 PT target nucleic acid in a biological sample obtained from the subject with
 PT a reagent.
 XX
 XX Claim 32; SEQ ID NO 366; 74pp; English.
 PS
 XX The invention relates to a novel method for detecting and differentiating
 CC between colon cell proliferative disorders associated with at least one
 CC gene or its regulatory regions. The method comprises contacting a target
 CC nucleic acid in a biological sample obtained from the subject with at
 CC least one reagent or a series of reagents, where the reagent or series of
 CC reagents, distinguishes between methylated and non methylated CpG
 CC dinucleotides within the target nucleic acid. The molecules of the
 CC invention demonstrate cytostatic activity whilst the method may useful
 CC for detecting and differentiating between colon cell proliferative
 CC disorders, including cancers such as colon adenoma and colon carcinoma.
 CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
 CC determining cytosine methylation state or single nucleotide
 CC polymorphisms. The current sequence is that of the pretreated genomic DNA
 CC region of the invention. This sequence is not shown within the
 CC specification but is taken from Wipoweb.
 XX
 SQ Sequence 5683 BP; 1357 A; 0 C; 1529 G; 2797 T; 0 U; 0 Other;

Query Match 8.1%; Score 41.2; DB 9; Length 5683;
 Best Local Similarity 56.7%; Pred. No. 0.12;
 Matches 76; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 365 TAACTGCTCCAGACCCCTACACCTCTAAGCAATCCTCTACTCAACTTATTCACAGAAC 424
 Db 1076 TAACTGCTCCAGACCCCTACACCTCTTTCAAAATACTTAACTACCCCAACCCACATAAC 1017
 QY 425 AAGCAATACAAACAGATGCCAAATCTCCACACTCTCGGACTACCCAAATTCGA 484
 Db 1016 ACAAATAACAAACACCAACCCATACACCAACCACTACATAAACAACCAACTTTT 957
 QY 485 CTGATATTTTGA 498
 Db 956 CTAAATAATCTAA 943

RESULT 11
 ACF62813/c
 ID ACF62813 standard; DNA; 5085 BP.

XX	AAEN56280;	RESULT 14	RAAS46793/C
XX	15-JUL-2002 (first entry)	ID	AAAS46793 standard; DNA; 56153 BP.
XX	Mouse spliced transcript detection oligonucleotide SEQ ID NO:29028.	XX	AAAS46793;
XX	Human; mouse; rat; splice transcript; detection; RNA transcript;	DT	18-DEC-2001 (first entry)
XX	splice variant; transcriptome; oligonucleotide library; ss.	DE	Tumour suppressor gene derived chemically modified sequence #519.
XX	Mus musculus.	XX	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
XX	WO200210449-A2.	KW	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX	07-FEB-2002.	KW	cytosine methylation; ds.
XX	20-JUL-2001; 2001WO-IB001903.	OS	Homo sapiens.
XX	28-JUL-2000; 2000US-0221607P.	XX	WO200168912-A2.
XX	02-MAY-2001; 2001US-0287724P.	PN	20-SEP-2001.
XX	(COMP-) COMPUGEN INC.	PD	15-MAR-2001; 2001WO-EP002955.
XX	Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;	PF	15-MAR-2000; 2000DE-01013847.
XX	WPI; 2002-257383/30.	PR	06-APR-2000; 2000DE-01019058.
XX	New oligonucleotide libraries comprising oligonucleotides which	PR	07-APR-2000; 2000DE-01019173.
XX	selectively hybridize to mRNAs transcribed from a transcription unit of a	PR	30-JUN-2000; 2000DE-01032529.
XX	genome, useful for detecting tissue-, pathology-, and developmental-	PR	01-SEP-2000; 2000DE-01043826.
XX	specific genes.	XX	(EPITG-) EPIGENOMICS AG.
XX	Example 1; SEQ ID NO 29028; 47pp; English.	PA	Olek A, Piepenbrock C, Berlin K;
XX	The present invention describes oligonucleotide libraries for detecting	PI	WPI; 2001-602752/68.
XX	messenger RNAs that populate a (sub-)transcriptome, where the (sub-	DR	Fragments of chemically modified genes associated with tumor suppressor
XX)transcriptome comprises messenger RNAs transcribed from multiple	XX	genes and oncogenes, useful in designing primers and probes for analyzing
XX	transcription units that populate a genome. The library comprises several	PT	diseases associated with cytosine methylation state e.g. cancer.
XX	oligonucleotides, each capable of hybridising selectively to a set of	XX	Claim 1; SEQ ID NO 519; 27pp; English.
XX	messenger RNAs transcribed from a given transcription unit of the genome,	XX	The invention relates to a nucleic acid comprising a sequence of 18
XX	which encodes one or more messenger RNA splice variants. The	CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX	oligonucleotide libraries are useful for detecting mRNAs from a	CC	bisulphite, of genes associated with tumour suppression and oncogenes
XX	biological sample, in expression profiling studies, in qualitatively or	CC	having a sequence taken from 536 (actually 533 since numbers 408, 458 and
XX	quantitatively characterising the corresponding transcriptome, and in	CC	500 are missing from the sequence listing) sequences (ss) and sequences
XX	detecting RNA transcripts and splice variants of human or animal	CC	complementary to (ss). The nucleic acid may be a peptide nucleic acid-
XX	transcriptomes. The libraries may also be used as specialised mini	CC	oligomer (PNA) of at least 9 nucleotides and may form part of a set of
XX	libraries to detect transcripts of a sub-transcriptome under a particular	CC	probes for detecting the cytosine methylation state and/or single
XX	biological or pathological state, and so allowing the detection of tissue	CC	nucleotide polymorphisms and also to be used in an array for analysing
XX	- and pathology-specific genes such as those genes only expressed in	CC	diseases associated with CpG dinucleotides e.g. cancers and tumours. The
XX	specific tissue under a specific pathological condition; to detect	CC	probes can also be used in a method for ascertaining genetic and/or
XX	developmental specific genes; and to detect RNA transcripts and splice	CC	epigenetic parameters for the diagnosis and/or therapy of existing
XX	variants of a transcriptome of a patient suffering from a particular	CC	diseases or the predisposition to specific diseases, by analysing
XX	disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from	CC	cytosine methylations. The parameters may be compared to another set of
XX	rats, humans and mice, which are used in the exemplification of the	CC	genetic and/or epigenetic parameters, which are disadvantageous to
XX	present invention. N.B. The sequence data for this patent did not form	CC	for diagnosis and/or prognosis events is one of the 533 genomic sequences
XX	part of the printed specification, but was obtained in electronic format	CC	patients. The present sequence is one of the 533 genomic sequences
XX	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	CC	derived from tumour suppressor genes and oncogenes. Note: The sequence
XX	Sequence 65 BP; 20 A; 21 C; 9 G; 15 T; 0 U; 0 Other;	CC	data for this patent did not form part of the printed specification, but
XX	Query Match 7.7%; Score 39.4; DB 6; Length 65;	CC	was obtained in electronic format directly from WIPO at
XX	Best Local Similarity 75.4%; Pred. No. 0.051;	CC	ftp.wipo.int/pub/published_pct_sequences
XX	Matches 49; Conservative 0; Mismatches 16; Indels 0; Gaps 0;	XX	Sequence 56153 BP; 13943 A; 1002 C; 14095 G; 27113 T; 0 U; 0 Other;
QY	424 CAAGCAAAATACAAAGAGATGCGAAATCTCCACACTACTGGACTACCCAAATTC 493	QY	Query Match 7.7%; Score 39.4; DB 4; Length 56153;
Db	1 CAAGCAACTACATGCGAGTGGCGAGTACTCTCCATACTACTCTACTGCTAGATTC 60	XX	Best Local Similarity 50.8%; Pred. No. 1.3;
QY	484 ACTGA 488	QY	Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
Db	61 ACTGA 65	QY	149 CAAGAGACAAATATCTTAGAGACACAAAGATCTCTCACTCTCTCTATCTTG 208
		Db	11877 CAAGCATCTAACAAAAATATACAGAAAAAACCCCTAAAAAAAACCGTAAACACTAT 11818
		QY	209 GTCTTCAGCTGATCCCAATGGTGGCAAAATAGGAGTAAACATCACTATACCTCTAAATC 268

This Page Blank (uspiu,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 15:55:56 ; Search time 2522 Seconds
(without alignments)
8764.848 Million cell updates/sec

Title: US-09-386-850-7_COPY_1_510

Perfect score: 510

Sequence: 1 AAAGTACTGACACAGAGAGC.....TTTTTGAAGTGTGGCAA 510

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.tod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	510	100.0	652	10	RATSMR1 M63112 Rat SMR1 pr
2	510	100.0	658	6	A07543 R.norvegicu
3	443	86.9	563	10	RNSMR1A1 X77819 R.norvegicu
4	389.8	76.4	641	10	RNSMR1A2 X77815 R.norvegicu
5	387	75.9	160624	2	AC144418 Rattus no
6	387	75.9	241010	2	AC103030 Rattus no
7	385.4	75.6	6433	10	RATSMR1A M59467 Rattus norv
8	385.4	75.6	6533	10	RNSMR1G X52467 R.norvegicu
9	385.4	75.6	8100	10	RNYCSA1 X84997 R.norvegicu
10	283.4	55.6	1618	10	RNSMR1A2G X77817 R.norvegicu
11	262.2	51.4	592	10	RNVCSA3 X77818 R.norvegicu
12	185.8	36.4	676	10	MMU82378 U82378 Mus musculu
13	179.4	35.2	676	10	MMU82377 U82377 Mus musculu
14	158.4	31.1	660	10	RNPRVBL X77816 R.norvegicu
15	155.6	30.5	680	10	MMU82380 U82380 Mus musculu
16	153	30.0	813	10	MMU82379 U82379 Mus musculu
17	151.6	29.7	743	10	MMU82376 U82376 Mus musculu
18	142	27.8	639	10	MMMSG1 X71629 M.musculus
19	142	27.8	653	10	BC034553 Mus muscu
20	142	27.8	671	10	BC055857 Mus muscu
21	139.6	27.4	655	10	BC031806 Mus muscu
22	124.2	24.4	205580	2	AC101839 Mus muscu
23	122.4	24.0	631	10	BC031921 Mus muscu
24	113.8	22.3	23003	10	MMU82375 U82375 Mus musculu
25	113.8	22.3	190544	2	AC138398 Mus musculu
26	113.8	22.3	205580	2	AC101839 Mus muscu
27	111	21.8	190544	2	AC138398 Mus muscu
28	104.8	20.5	471	10	MMMSG3 X71631 M.musculus
29	89.2	17.5	84611	2	AC139235 Mus muscu
30	88.4	17.3	61021	2	AC101916 Mus muscu
31	76.8	15.1	8239	10	RNVCS X74229 R.rattus VC
32	76.8	15.1	257158	2	AC122599 Rattus no
33	70.6	13.8	555	10	BC059094 Mus muscu
34	63.4	12.4	160624	2	AC144418 Rattus no
35	57.4	11.3	61021	2	AC101916 Mus muscu
36	56.8	11.1	366	6	BD071680 Secretd
37	56.8	11.1	521	6	AX892590 Sequence
38	56.8	11.1	521	6	BD028123 Sequence
39	56.8	11.1	723	11	G28337 human STS S
40	56.8	11.1	724	9	HUMRRP D29833 Homo sapien
41	56.8	11.1	733	9	BC015327 Homo sapi
42	56.8	11.1	957	9	AK130485 Homo sapi
43	53.4	10.5	153	6	AX381686 Sequence
44	51	10.0	84611	2	AC139235 Mus muscu
45	47.4	9.3	195311	10	AC121585 Mus muscu

ALIGNMENTS

RESULT 1
RATSMR1
LOCUS
DEFINITION
Rat SMR1 protein mRNA, complete cds.
ACCESSION
M63112 J04109
VERSION
M63112.1 GI:206997
KEYWORDS
SMR1 protein.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 652)
AUTHORS
Rosinski-Chupin,I., Tronik,D. and Rougeon,F.
TITLE
High level of accumulation of a mRNA coding for a precursor-like

JOURNAL protein in the submaxillary gland of male rats
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 85 (22), 8553-8557 (1988)
PUBMED 89042220
COMMENT 3186744
Original source text: Rat (Wistar, male) submaxillary gland, cDNA
to mRNA, clone CD 13.
Draft entry and computer-readable sequence for [1] kindly provided
by I Rosinski-Chupin, 14-NOV-1988.

FEATURES

Location/Qualifiers
1..652
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
c1..652
/product="SMG1 mRNA"
73..513
/note="SMG1 protein precursor"
/codon_start=1
/protein_id="AAA42153.1"
/db_xref="GI:206998"
/translation="MKSLYLIFGLMILLACFQSGVGRGPRRQHPRRQDDPSTLPHY
LGLPDPNGGQIGVTITPLNLPQPRVLNLPFGITGPPVVGTTGYOYQWLTPAD
PTPLSNPPTQLHSTEQANTKDAKISNTTATTQNSTDIFEGGK"
sig_peptide 73..126
/note="SMG1 protein signal peptide"
mat_peptide 127..510
/product="SMG1 protein"

ORIGIN

Query Match 100.0%; Score 510; DB 10; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.8e-134;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAATGACTGACGAGAGCTTCTGACGACACATTTCCCGCTCAGAGTTTCTCCAAG 60
Db 1 AAATGACTGACGAGAGCTTCTGACGACACATTTCCCGCTCAGAGTTTCTCCAAG 60
QY 61 GGGCTACCAAGATGAAGTCACTGTATTTGATCTTTGGCTGTGGATCTTCTAGCATGC 120
Db 61 GGGCTACCAAGATGAAGTCACTGTATTTGATCTTTGGCTGTGGATCTTCTAGCATGC 120
QY 121 TTCCAGTCAGGTGAGGCTGTGAGGCGCCCAAGAGACAACTAATCCTAGAGACAA 180
Db 121 TTCCAGTCAGGTGAGGCTGTGAGGCGCCCAAGAGACAACTAATCCTAGAGACAA 180
QY 181 GATCCTTCAACTCTTCTCAATATCTTGGTCTTTCAGCCTGATCCCAATGGTGACAA 240
Db 181 GATCCTTCAACTCTTCTCAATATCTTGGTCTTTCAGCCTGATCCCAATGGTGACAA 240
QY 241 GGAGTAACATCACTATACCTTAAATCTTCAACCACTCGTCTTCTTGTAAATCTTCCC 300
Db 241 GGAGTAACATCACTATACCTTAAATCTTCAACCACTCGTCTTCTTGTAAATCTTCCC 300
QY 301 GGTITTTATCACTGGACCACTGGTTGTACAGGTACCACTGAATATCAATATCAGTGG 360
Db 301 GGTITTTATCACTGGACCACTGGTTGTACAGGTACCACTGAATATCAATATCAGTGG 360
QY 361 CAGCTAATCTGCTCAGACCCCTACACCTCTTAAGCAATCTCTCTACTCACTTCCACA 420
Db 361 CAGCTAATCTGCTCAGACCCCTACACCTCTTAAGCAATCTCTCTACTCACTTCCACA 420
QY 421 GAACAGCAATACAAACAGATGCCAAATCTCCAACTACTGGACTACCCAAAT 480
Db 421 GAACAGCAATACAAACAGATGCCAAATCTCCAACTACTGGACTACCCAAAT 480
QY 481 TCCACTGATATTTTGAAGTGGTGCAAA 510
Db 481 TCCACTGATATTTTGAAGTGGTGCAAA 510

RESULT 2

A07543 658 bp RNA linear PAT 27-JUL-1993
LOCUS

DEFINITION

R.norvegicus SMR1 mRNA.

ACCESSION

A07543 GI:412272

VERSION

SMR1 protein.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 658)

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches 510;

Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

```

RESULT 3
RNSMRL1A1
LOCUS      563 bp      mRNA      linear      ROD 10-MAR-2001
DEFINITION R.norvegicus VCS-alpha1 mRNA for SMR1-alpha1.
ACCESSION  X77819
VERSION     X77819.1 GI:732921
KEYWORDS   SMR1 gene.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1
            Singer, M., Courty, Y., and Rougeon, F.
            Recent evolution of genes encoding the prohormone-like protein SMR1
            in the rat submandibular gland
            DNA Cell Biol. 14 (2), 137-144 (1995)
JOURNAL    95169272
MEDLINE    7865131
PUBMED
REFERENCE  2
            Courty, Y., Singer, M., Rosinski-Chupin, I., and Rougeon, F.
            Episodic evolution and rapid divergence of members of the rat
            multigene family encoding the salivary prohormone-like protein SMR1
            Mol. Biol. Evol. 13 (6), 758-766 (1996)
JOURNAL    96276306
MEDLINE
PUBMED
REFERENCE  3
            Courty, Y., Singer, M., Rosinski-Chupin, I., and Rougeon, F.
            A new proline-rich protein precursor expressed in the salivary
            glands of the rat is encoded by a gene homologous to the gene
            coding for the prohormone-like protein SMR1
            Unpublished
            Location/Qualifiers
            1..563
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /clone="submandibular gland RT-PCR"
            /tissue_type="submandibular gland"
            /dev_stage="adult"
            1..563
            /gene="VCS-alpha1"
            46..486
            /gene="VCS-alpha1"
            /codon_start=1
            /product="SMR1-alpha1"
            /protein_id="CAA54834.1"
            /db_xref="GI:732922"
            /db_xref="SPTREMBL:Q63552"
            /translation="MKPLYLFIQLILIGCFOSGGRGPRFOHNLRRQTPSTLRHY
            LGLQDPNDVQMRVTITPLNLQPPRLVNLPGFITGPLVVGQSTEQYQQLTAPD
            PRPLSHPTQLHSTEQANTKTDKLSNTTANTQNSTDIFEGGK"
            112..483
            /gene="VCS-alpha1"
            /product="SMR1-alpha1"

mat_peptide
            86.9%; Score 443; DB 10; Length 563;
            Best Local Similarity 94.8%; Pred. No. 2.4e-115;
            Matches 458; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY      28  CAGCACATTTCCCGCTCAGAAGTTTCTCCAAAGGGGCTACCAAGATGAAGTCACTGTAT 87
Db      1  CAGCACATTTCCCGCTCAGAAGTTTCTCCAAAGGGGCTACCAAGATGAAGTCACTGTAT 60

QY      88  TTGATCTTTGGCCTGTGGATCTCTTAGCATCTTCCAGTCAAGGTGAGGGTGTACAGAGC 147
Db      61  TTGATCTTTGGCCTGTGGATCTTATAGGATCTTCCAGTCAAGGGAGGGTACCGAGGC 120

QY      148  CCAAGAGACACATATCTTAGAGACACACAAAGATCCCTTCAACTCTTCCCTCATTTATCTT 207
Db      121  CCGAAGAGACACATAATCTTAGAGAGCGACAACTCCTTCAACTCTTCCCTCATTTATCTT 180

```

```

QY      208  GGCTTCAGCCTGATCCCAATGGTGGCAAAATAGGAGTAACAATCACTATACCTTTAAAT 267
Db      181  GGCTTCAGCCTGATCCCAATGATGTGCAATAGAGAGTAACAATCACTATACCTTTAAAT 240

QY      268  CTTCAACACCACTCGTCTCTTCTTAACTCTCCCGGTTTATCACTGAGACCACTTGGTT 327
Db      241  CTTCAACACCACTCGTCTCTTCTTAACTCTCCCGGTTTATCACTGAGACCACTTGGTT 300

QY      328  GTACAAGGTACCACTGAATATCAATATCACTGAGTGGCAGCTAACTGCTCCAGACCTTACACCT 387
Db      301  GTACAAGGTGCCACAGAATATCAATATCACTGAGTGGCAGCTAACTGCTCCAGACCTTACACCT 360

QY      388  CTAAGCAATCTCTTACTCACTTCACTTCCACAGAACAGCAAGCAATAACAAAACAGATGCC 447
Db      361  CTAAGCAATCTCTTACTCACTTCACTTCCACAGAACAGCAAGCAATAACAAAACAGATGCC 420

QY      448  AAAATCTCCAACACTACTCGCTACTCCCAAAATTCACATGATATTTTGAAGTGGTGGC 507
Db      421  AAAATCTCCAACACTACTCGCTACTCCCAAAATTCACATGATATTTTGAAGTGGTGGC 480

QY      508  AAA 510
Db      481  AAA 483

RESULT 4
RNSMRL1A2
LOCUS      641 bp      mRNA      linear      ROD 10-MAR-2001
DEFINITION R.norvegicus VCS-alpha2 mRNA for SMR1-alpha2.
ACCESSION  X77815
VERSION     X77815.1 GI:1563744
KEYWORDS   SMR1 gene.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1
            Courty, Y., Singer, M., Rosinski-Chupin, I., and Rougeon, F.
            Episodic evolution and rapid divergence of members of the rat
            multigene family encoding the salivary prohormone-like protein SMR1
            Mol. Biol. Evol. 13 (6), 758-766 (1996)
JOURNAL    96276306
MEDLINE    8754212
PUBMED
REFERENCE  2
            Courty, Y., Rosinski-Chupin, I., and Rougeon, F.
            A new proline-rich protein precursor expressed in the salivary
            glands of the rat is encoded by a gene homologous to the gene
            coding for the prohormone-like protein SMR1
            Unpublished
            Location/Qualifiers
            1..641
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /strain="Wistar"
            /db_xref="taxon:10116"
            /clone="ZAP33"
            /tissue_type="submandibular gland"
            /dev_stage="immature rat (18 days old)"
            1..641
            /gene="VCS-alpha2"
            join(1..53,54..119,120..641)
            /gene="VCS-alpha2"
            1..53
            /gene="VCS-alpha2"
            /number=1
            54..119
            /gene="VCS-alpha2"
            /number=2
            join(66..119,120..413)
            /gene="VCS-alpha2"
            /codon_start=1
            /product="SMR1-alpha2"

FEATURES
            source
            1..641

```


* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 46 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	1	1226:	contig of 1226 bp in length
*	1227	1326:	gap of unknown length
*	1327	2508:	contig of 1182 bp in length
*	2509	2608:	gap of unknown length
*	2609	4039:	contig of 1431 bp in length
*	4040	4139:	gap of unknown length
*	4140	5356:	contig of 1217 bp in length
*	5357	5456:	gap of unknown length
*	5457	6529:	contig of 1073 bp in length
*	6530	6629:	gap of unknown length
*	6630	7800:	contig of 1171 bp in length
*	7801	7900:	gap of unknown length
*	7901	9544:	contig of 1644 bp in length
*	9545	9644:	gap of unknown length
*	9645	10817:	contig of 1173 bp in length
*	10818	10917:	gap of unknown length
*	10919	12521:	contig of 1604 bp in length
*	12522	12621:	gap of unknown length
*	12622	14406:	contig of 1785 bp in length
*	14407	14506:	gap of unknown length
*	14507	16291:	contig of 1785 bp in length
*	16292	16391:	gap of unknown length
*	16392	17664:	contig of 1273 bp in length
*	17665	17764:	gap of unknown length
*	17765	19397:	contig of 1633 bp in length
*	19398	19497:	gap of unknown length
*	19498	21655:	contig of 2158 bp in length
*	21656	21755:	gap of unknown length
*	21756	22961:	contig of 1206 bp in length
*	22962	23061:	gap of unknown length
*	23062	24831:	contig of 1770 bp in length
*	24832	24931:	gap of unknown length
*	24932	26934:	contig of 2003 bp in length
*	26935	27034:	gap of unknown length
*	27035	29028:	contig of 1994 bp in length
*	29029	29128:	gap of unknown length
*	29129	31188:	contig of 2060 bp in length
*	31189	31288:	gap of unknown length
*	31289	32978:	contig of 1690 bp in length
*	32979	33078:	gap of unknown length
*	33079	34291:	contig of 1213 bp in length
*	34292	34391:	gap of unknown length
*	34392	36429:	contig of 2038 bp in length
*	36430	36529:	gap of unknown length
*	36530	38982:	contig of 2453 bp in length
*	38983	39082:	gap of unknown length
*	39083	40236:	contig of 1154 bp in length
*	40237	40336:	gap of unknown length
*	40337	42653:	contig of 2317 bp in length
*	42654	42753:	gap of unknown length
*	42754	44497:	contig of 1744 bp in length
*	44498	44597:	gap of unknown length
*	44598	47423:	contig of 2826 bp in length
*	47424	47523:	gap of unknown length
*	47524	48871:	contig of 1348 bp in length
*	48872	48971:	gap of unknown length
*	48972	51424:	contig of 2453 bp in length
*	51425	51524:	gap of unknown length
*	51525	55311:	contig of 3787 bp in length
*	55312	55411:	gap of unknown length
*	55412	59896:	contig of 4485 bp in length
*	59897	63366:	contig of 3370 bp in length
*	63367	63466:	gap of unknown length

63467	66715:	contig	of 3249 bp in length
66716	66815:	gap	of unknown length
66916	66965:	contig	of 2880 bp in length
66966	67975:	gap	of unknown length
67976	73512:	contig	of 3717 bp in length
73513	73612:	gap	of unknown length
73613	77665:	contig	of 4053 bp in length
77666	82977:	gap	of unknown length
82978	83077:	contig	of 5212 bp in length
83078	86589:	contig	of 3512 bp in length
86590	86689:	gap	of unknown length
86690	91066:	contig	of 4377 bp in length
91067	91166:	gap	of unknown length
91167	92290:	contig	of 1124 bp in length
92291	92390:	gap	of unknown length
92391	97723:	contig	of 5332 bp in length
97723	97823:	gap	of unknown length
97823	104887:	contig	of 7065 bp in length
104888	104987:	gap	of unknown length
104988	111110:	contig	of 6123 bp in length
111111	111110:	gap	of unknown length
112121	123123:	contig	of 11919 bp in length
123123	123228:	gap	of unknown length
123230	138677:	contig	of 13448 bp in length
138678	138777:	gap	of unknown length
138778	1506324:	contig	of 21847 bp in length.

FEATURES

```

source
1. .160624
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-366p1"

```

ORIGIN

Query Match 75.9%; Score 387; DB 2; Length 150624;
Best Local Similarity 100.0%; Pred. No. 3.2e-99;
Matches 387; Conservative 0; Mismatches 0; Indels 0;

124	CAGT	CAGT	GAGGGTGT	CTAGAGCCCAAGACAA	CATAATCTTAGAAGACAA	CAAGAT	183	
139022	CAGT	CAGT	GAGGGTGT	CTAGAGCCCAAGACAA	CATAATCTTAGAAGACAA	CAAGAT	139081	
184	CCTT	CAACT	CTCTCCAT	TATCTTGGTCTTCAGCC	TGATCCCAATGGTGGCA	AAATAGGA	243	
139082	CCTT	CAACT	CTCTCCAT	TATCTTGGTCTTCAGCC	TGATCCCAATGGTGGCA	AAATAGGA	139141	
244	GTA	CAACAT	CAC	TATACCC	TAAATCTTCAAC	CAACCTCGTGTCTTGT	TAAATCTTCCCGT	303
139142	GTA	CAACAT	CAC	TATACCC	TAAATCTTCAAC	CAACCTCGTGTCTTGT	TAAATCTTCCCGT	139201
304	TTT	TAT	CAT	CTGGAC	CACCAATGGTGT	TACAAGGTACCA	CTGAATCAATATCAGTGGCAG	363
139202	TTT	TAT	CAT	CTGGAC	CACCAATGGTGT	TACAAGGTACCA	CTGAATCAATATCAGTGGCAG	139261
364	CTA	CTGCT	CCAGAC	CCCTAC	ACCTCTAAGCAAT	CTCTCTACTCAACT	TTCATTCCACAGAA	423
139262	CTA	CTGCT	CCAGAC	CCCTAC	ACCTCTAAGCAAT	CTCTCTACTCAACT	TTCATTCCACAGAA	139321
424	CAAG	CAAA	TACAAA	AACAGATG	CCAAAATCTTCCAAC	ACTACTGGGAC	TACCCAAAATTC	483
139322	CAAG	CAAA	TACAAA	AACAGATG	CCAAAATCTTCCAAC	ACTACTGGGAC	TACCCAAAATTC	139381
484	ACT	GAT	ATTTT	TGAAGT	GTGGCAAA	510		
139382	ACT	GAT	ATTTT	TGAAGT	GTGGCAAA	139408		

RESULT 6

AC103030/c

LOCUS

LOCUS	AC103030	241010 bp	DNA linear	HTG 13-MAY-2003
DEFINITION	<i>Rattus norvegicus</i>	clone CH230-203K20,	WORKING DRAFT	SEQUENCE, 3
NOTES	unordered pieces.			

COMMENT

CDS	gene="VCSA1" Join(2235..2288,6844..6996) /gene="VCSA1" /codon_start=1 /product="P2-Va1" /protein_id="CAA59356.1" /db_xref="GI:693998" /db_xref="SPTREMBL:Q63671" /translation="MKSLYLIFGLWILLACFQTAGMIGVCLPTPLNFSKIFLKVSQFV" PCHISPTVTVRVPYSCDNFLTP" Join(2235..2288,5604..5990) /gene="VCSA1" /codon_start=1 /product="SMR1-Va1" /protein_id="CAA59355.1" /db_xref="GI:693997" /db_xref="SWISS-PROT:P13432" /translation="MKSLYLIFGLWILLACFQSGEGVGRPRRHNPRRODPSTPLHY" LGLPDNGQIGVITITPLNPGFVRLVNLPGFIVPLVVGQITTEYQVQWOLTAPD" PTPLSNPTTOLLSEQANTTKDAKISNTTATTONSTDIFEGGK" 2289..5603 /gene="VCSA1" /number=2 /evidence=experimental 5604..6130 /gene="VCSA1" /number=3 /evidence=experimental 5890 /gene="VCSA1" /replace="a" 6131..6843 /gene="VCSA1" /number=3 /evidence=experimental 6350..6354 /gene="VCSA1" /replace="" 6374 /gene="VCSA1" /replace="a" 6607 /gene="VCSA1" /replace="c" 6657 /gene="VCSA1" /replace="c" 6670 /gene="VCSA1" /replace="c" 6844..7403 /number=4 /evidence=experimental	
CDS		
intron		
exon		
variation		
intron		
variation		
variation		
variation		
variation		
exon		
ORIGIN		
Query Match	75.6%; Score 385.4; DB 10; Length 8100;	
Best Local Similarity	99.7%; Pred. No. 7.7e-99;	
Matches 386; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	124 CAGTCAGGTGAGGGTCTCAGAGGCCCAAGACACATATAATCCTAGAAGACAAAGAT	183
Db	5601 CAGTCAGGTGAGGGTCTCAGAGGCCCAAGACACATATAATCCTAGAAGACAAAGAT	5660
QY	184 CCTTCAACTCTTCCTCATATCTTGTGTTTCAGCCGTATCCCAATGGTGACAAATAGA	243
Db	5661 CCTTCAACTCTTCCTCATATCTTGTGTTTCAGCCGTATCCCAATGGTGACAAATAGA	5720
QY	244 GTAAACAATCACTATACCCCTTAATATCTTCAACCACCTCGTGTGTTCTTGTTAATCTTCCCGGT	303
Db	5721 GTAAACAATCACTATACCCCTTAATATCTTCAACCACCTCGTGTGTTCTTGTTAATCTTCCCGGT	5780
QY	304 TTATATCACTGGACCAACATGGTGTGTAAGGTACCACTGAATATCAATATCACTGGGACG	363

Query Match 55.6%; Score 283.4; DB 10; Length 1618;
 Best Local Similarity 85.3%; Pred. No. 9.4e-70;
 Matches 330; Conservative 0; Mismatches 51; Indels 6; Gaps 1;
 Qy 124 CAGTCAGGTGAGGTGTCAGAGGCCAGAGAGACACATATCTCTAGAGACACACAGAT 183
 Db 1078 CAGTCAGGTGAGGTGTCAGAGGCCAGAGAGACACATATCTCTAGAGACACACAGAT 1137
 Qy 184 CTTCAACTCTTCTCATATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATAGGA 243
 Db 1138 CTTCAACTCTTCTCGTATTTCTGCTTCAGTCTGATCTTAATGGTGGACAAATAGGA 1197
 Qy 244 GTAAACAATCATTACCTTTAAATCTTCAACACCTCGTGTCTTTTAAATCTTCCCGT 303
 Db 1198 GAAACAACAGTATACCTTTAACTATTTCAACACCTAGCAGAAATTTATGATCTTCTCTGA 1257
 Qy 304 TTTATCATGGACCAACCAATCTGTTGTACAGGTACCACTGAATATCAATATCAGTGGCAG 363
 Db 1258 ATTITTAACATAAATAACAGTGGTCACTAGTACCATTTG-----AATATCATTTGTGAG 1311
 Qy 364 CTAACCTGCTCCAGACCTTACCTCTAAGCAATCTCTCTACTCAACTTCAATTCACAGAA 423
 Db 1312 CTAACCTGCTCCAGACCTTCTAAGCAATCTCTCTACTCAACTTCAATTCACAGAA 1371
 Qy 424 CAAGCAATATACAAAACAGATGCCAAATCTCCACACTACTGCGACTACCCAAATTC 483
 Db 1372 TAAGCAAAATACAAAACAGATGCCAAATCTCCACACTACTGCGACTACCCAAATTC 1431
 Qy 484 ACTGATATTTTGAAGGTGGTGGCAA 510
 Db 1432 ACTGATATTTTGAAGGTGGTGGCAA 1458

RESULT 11
 RNVC5A3 592 bp DNA linear ROD 10-MAR-2001
 LOCUS R.norvegicus VCS-alpha3 gene.
 DEFINITION X77818
 ACCESSION X77818.1 GI:1563748
 VERSION SMR1 gene.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1
 AUTHORS Courtney, Y., Singer, M., Rosinski-Chupin, I. and Rougeon, F.
 TITLE Episodic evolution and rapid divergence of members of the rat
 multigene family encoding the salivary prohormone-like protein SMR1
 JOURNAL Mol. Biol. Evol. 13 (6), 758-766 (1996)
 MEDLINE 96276306
 PUBMED 8754212
 REFERENCE 2 (bases 1 to 592)
 AUTHORS Courtney, Y., Rosinski-Chupin, I. and Rougeon, F.
 TITLE A new proline-rich protein precursor expressed in the salivary
 glands of the rat is encoded by a gene homologous to the gene
 coding for the prohormone-like protein SMR1
 JOURNAL Unpublished
 FEATURES Location/Qualifiers
 source 1. .592
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /clone="genomic PCR"
 /tissue type="spleen"
 135. .592
 /gene="VCS-alpha3"
 135. .592
 /gene="VCS-alpha3"
 /number=3
 <135. .500
 CDS

/gene="VCS-alpha3"
 /codon_start=1
 /product="SMR1-alpha3"
 /protein_id="CA54833.1"
 /db_xref="GI:1563748"
 /db_xref="SPTREMBL:P70634"
 /translation="GDRGQRHNPQRQDPSTLPHLCLQPNVQVQMLRNLKPSTI
 TVPPPLVCIPIPGFISGAVFNQLGTEHYVQCOLTAPDRSLSPNPQLHTTEOANTK
 TDDKISNTATTQNSITDF"
 ORIGIN
 Query Match 51.4%; Score 262.2; DB 10; Length 592;
 Best Local Similarity 82.9%; Pred. No. 1e-63;
 Matches 324; Conservative 0; Mismatches 63; Indels 4; Gaps 2;
 Qy 124 CAGTCAGGTGAGGTGTCAGAGGCCAGAGAGACACATATCTCTAGAGACACACAGAT 183
 Db 123 CAGTAAGGTGAGGTGTCAGAGGCCAGAGAGACACATATCTCTAGAGACACACAGAT 182
 Qy 184 CTTCAACTCTTCTCATATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATA 240
 Db 183 CTTCAACTCTTCTCATATCTTGGTCTTCAGCTGATCCCAATGGTGGACAAATA 242
 Qy 241 GGAGTAAACAATCATTACCTTTAAATCTTCAACACCTCGTGTCTTTTAAATCTTCC 300
 Db 243 AGAAATTTACCCCTCTAGTACATATACAGTTGATCCACCTCTTTCTTTGTTGATCTCT 302
 Qy 301 GGTTTTACTCTGGACCACTTGGTGTCAAGGTACACCTGAAATATCAATATCAGTGG 360
 Db 303 GGTTTTATCAGTGGAGCAGTATTCATCACTAGTACACCTGATATCATATCAGTGT 362
 Qy 361 CAGCTAACTGCTCCAGACCTACACCTCTAAGCAATCTCTCTACTCACTTCACTCCACA 420
 Db 363 CAGCTAACTGCTCCAGACCTAGATCTCTAAGCAATCTCTCTAATCACTTCACTACCACA 422
 Qy 421 GAACAGCAATACAAAACAGATGCCAAATCTCCACACTACTGCGACTACCCAAAT 480
 Db 423 GAACAGCAATACAAAACAGATGCCAAATCTCCACACTACTGCGACTACCCAAAT 482
 Qy 481 TCCACTGATA-TTTTGAAGGTGGTGGCAA 510
 Db 483 TCCACTGATACTTTTGAAGGTGGTGGCAA 513

RESULT 12
 RNVC5A3 676 bp mRNA linear ROD 31-DEC-1997
 LOCUS Mus musculus MSG2delta salivary protein (Vcs2) mRNA, complete cds.
 DEFINITION U82378
 ACCESSION U82378
 VERSION U82378.1 GI:2289108
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 676)
 AUTHORS Senorale-Pose, M. and Rougeon, F.
 TITLE The mouse Vcs2 gene is a composite structure which evolved by gene
 fusion and encodes five distinct salivary mRNA species
 JOURNAL Gene 201 (1-2), 75-85 (1997)
 MEDLINE 98072432
 PUBMED 9409774
 REFERENCE 2 (bases 1 to 676)
 AUTHORS Senorale-Pose, M. and Rougeon, F.
 TITLE Direct Submission
 JOURNAL Submitted (16-DEC-1996) Immunologie, Institut Pasteur, 25 rue du
 Dr. Roux, Paris 75015, France
 FEATURES Location/Qualifiers
 source 1. .676
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"

gene /sex="male"
/tissue_type="submandibular gland"
1. .676
/gene="Vcs2"
88. .519
/gene="Vcs2"
/codon_start=1
/product="MSG2delta salivary protein"
/protein_id="AAB93516.1"
/db_xref="GI:2289109"
/translation="MKPLCLVGLCVLIGCFLLSSCQRCFRGQHDPTRLSPNSPSSH
FYQDPDRVQISQEDNIPFMPEQHSNLCVPPPLYLGEFEKLPENTHIPYILI
RPDIIPPYSKYIQVPRKSNATPAANNFTTATAPNSTDSF"

ORIGIN

Query Match 36.4%; Score 185.8; DB 10; Length 676;
Best Local Similarity 65.3%; Pred. No. 6.5e-42;
Matches 340; Conservative 0; Mismatches 167; Indels 14; Gaps 4;

QY 1 AAATGACTGACGAGAGCTTCTGACGACGACATTTCCCGCTCAGAAGTTTCTCCAAG 60
DB 15 AAAATAATTGACAGACTCTTCTGACGACGACATTCACGAGCAAGAAGCTTCTCCAAG 74
QY 61 GGGCTAC-CAAGATGAAGTCACTGTATTGTATCTTTGGCTCTGGATCCTTCTAGCATG 119
DB 75 AGGCTACTGAAAGATGAAGCACTGTGTTGGTCTTTGGCTCTGTGTTCTTATAGATG 134
QY 120 CTTCCAGTCAGGTGAGGTGTGAGAGGCCCAAGAGCAACATAATCTTAGAGACAACA 179
DB 135 CTTCCCTTTCTAGTAGTGCCAAAGAGGCTTCAGAGGACAGCATGATCTTACAAGACCCT 194
QY 180 AGATCCTTCAACTCTTCCCTCAATATCTTGGTCTTTCAGCTGATCCCAATGGTGGACAAT 239
DB 195 ATCTCCTTCAAAATCTTCTTCAATTTTATCTCAGCTGATCCCAATGGTGGACAAT 254
QY 240 AGGAGTAAACAATCACTATACCC-----TTAAATCTTCAACACCTCGTCTTCTGT 290
DB 255 TTCTCAACCAGACAATATCCCATTTTCATGTTGACAAACACATCTCTTAATATTG 314
QY 291 TAACTCTCCGGTTTATACCTGGACCAACATTTGGTTTACAGGTACCACTGAATATCA 350
DB 315 TGTTCCTCTCTCTCTTTATACCTTGGAGAGAATTTGAAAACACTCTCTTAATCTCA 374
QY 351 ATATCAGTGGCAGCTAACTGCTCCAGACCCCTACACCTCTTAAGCAATCTCTTACTCACT 410
DB 375 TATTCCTTATATATCTAATACGTCAGATATTGAACCTCCAAGCAA---ATATATTCAACC 431
QY 411 TCATTCCACAGAACAGCAAAATACAAAACAGATGCCAAATCTCCAACACTACTGGCAG 470
DB 432 AGTTCACAGAAAAAATCAATGCAACACCACTGCGCAATCTTATTAATCTACTGCTAC 491
QY 471 TACCCAAATTTCACTGATA-TTTTGAAGGTGGGCAAA 510
DB 492 TGCCCCCAATTTACTGATAGCTTTTGAAGGTGGTGGCAAA 532

RESULT 13
NMU82377
LOCUS
DEFINITION Mus musculus MSG2gamma salivary protein (Vcs2) mRNA, complete cds.
ACCESSION U82377
VERSION U82377.1 GI:2289106
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 676)
Senorale-Pose, M. and Rougeon, F.
The mouse Vcs2 gene is a composite structure which evolved by gene
fusion and encodes five distinct salivary mRNA species
JOURNAL Gene 201 (1-2), 75-85 (1997)
MEDLINE 98072432

9409774
REFERENCE 2 (bases 1 to 676)
AUTHORS Senorale-Pose, M. and Rougeon, F.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1996) Immunologie, Institut Pasteur, 25 rue du
Dr. Roux, Paris 75015, France
FEATURES
Location/Qualifiers
1. .676
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="submandibular gland"
1. .676
/gene="Vcs2"
88. .519
/gene="Vcs2"
/codon_start=1
/product="MSG2gamma salivary protein"
/protein_id="AAB93515.1"
/db_xref="GI:2289107"
/translation="MKALYMPVFLVULIGCFLLSSCQRCFRGQHDPTRLSPNSPSSH
FYQDPDRVQISQEDNIPFMPEQHSNLCVPPPLYLGEFEKLPENTHIPYILI
RPDIIPPYSKYIQVPRKSNATPAANNFTTATAPNSTDSF"

ORIGIN

Query Match 35.2%; Score 179.4; DB 10; Length 676;
Best Local Similarity 64.5%; Pred. No. 4.4e-40;
Matches 336; Conservative 0; Mismatches 171; Indels 14; Gaps 4;

QY 1 AAATGACTGACGAGAGCTTCTGACGACGACATTTCCCGCTCAGAAGTTTCTCCAAG 60
DB 15 AAAATAATTGACAGACTCTTCTGACGACGACATTCACGAGCAAGAAGCTTCTCCAAG 74
QY 61 GGGCTACCAA--GATGAAGTCACTGTATTGTATCTTTGGCTCTGGATCCTTCTAGCATG 119
DB 75 ATCTTAAAAAAGCATGAAAGGCACTGTATATGTTATTTGCTCTGGTCTTCTTATAGATG 134
QY 120 CTTCCAGTCAGGTGAGGTGTGAGAGGCCCAAGAGACAACATAATCTTAGAGACAACA 179
DB 135 TTCTCTGCTAGTAGTGCCAAAGAGGCTTCAGAGGACAGCATGATCTTACAAGACCCT 194
QY 180 AGATCCTTCAACTCTTCTCTCAATTTATTTGGTCTTTCAGCTGATCCCAATGGTGGACAAT 239
DB 195 ATCTCCTTCAAAATCTTCTTCAATTTTATCTCAGCTGATCCCAATGGTGGACAAT 254
QY 240 AGGAGTAAACAATCACTATACCC-----TTAAATCTTCAACACCTCGTCTTCTGT 290
DB 255 TTCTCAACAGACAATATCCCATTTTCATGTTTGAACAACACATCTCTTAAATATTG 314
QY 291 TAACTCTCCGGTTTATACCTGGACCAACATTTGGTTTACAGGTACCACTGAATATCA 350
DB 315 TGTTCCTCTCTCTCTTTATACCTTGGAGAGAATTTGAAAACACTCTCTTAATCTCA 374
QY 351 ATATCAGTGGCAGCTAACTGCTCCAGACCCCTACACCTCTTAAGCAATCTCTTACTCACT 410
DB 375 TATTCCTTATATATCTAATACGTCAGATATTGAACCTCCAAGCAA---ATATATTCAACC 431
QY 411 TCATTCCACAGAACAGCAAAATACAAAACAGATGCCAAATCTCCAACACTACTGGCAG 470
DB 432 AGTTCACAGAAAAAATCAATGCAACACCACTGCGCAATCTTATTAATCTACTGCTAC 491
QY 471 TACCCAAATTTCACTGATA-TTTTGAAGGTGGGCAAA 510
DB 492 TGCCCCCAATTTACTGATAGCTTTTGAAGGTGGTGGCAAA 532

RESULT 14
RNPVRB1
LOCUS
DEFINITION R. norvegicus VSC-beta1 mRNA for PR-Vbeta1.
ACCESSION X77816
RNPVRB1 660 bp mRNA linear ROD 11-MAR-2001

[illegible]

This Page Blank (uspto)